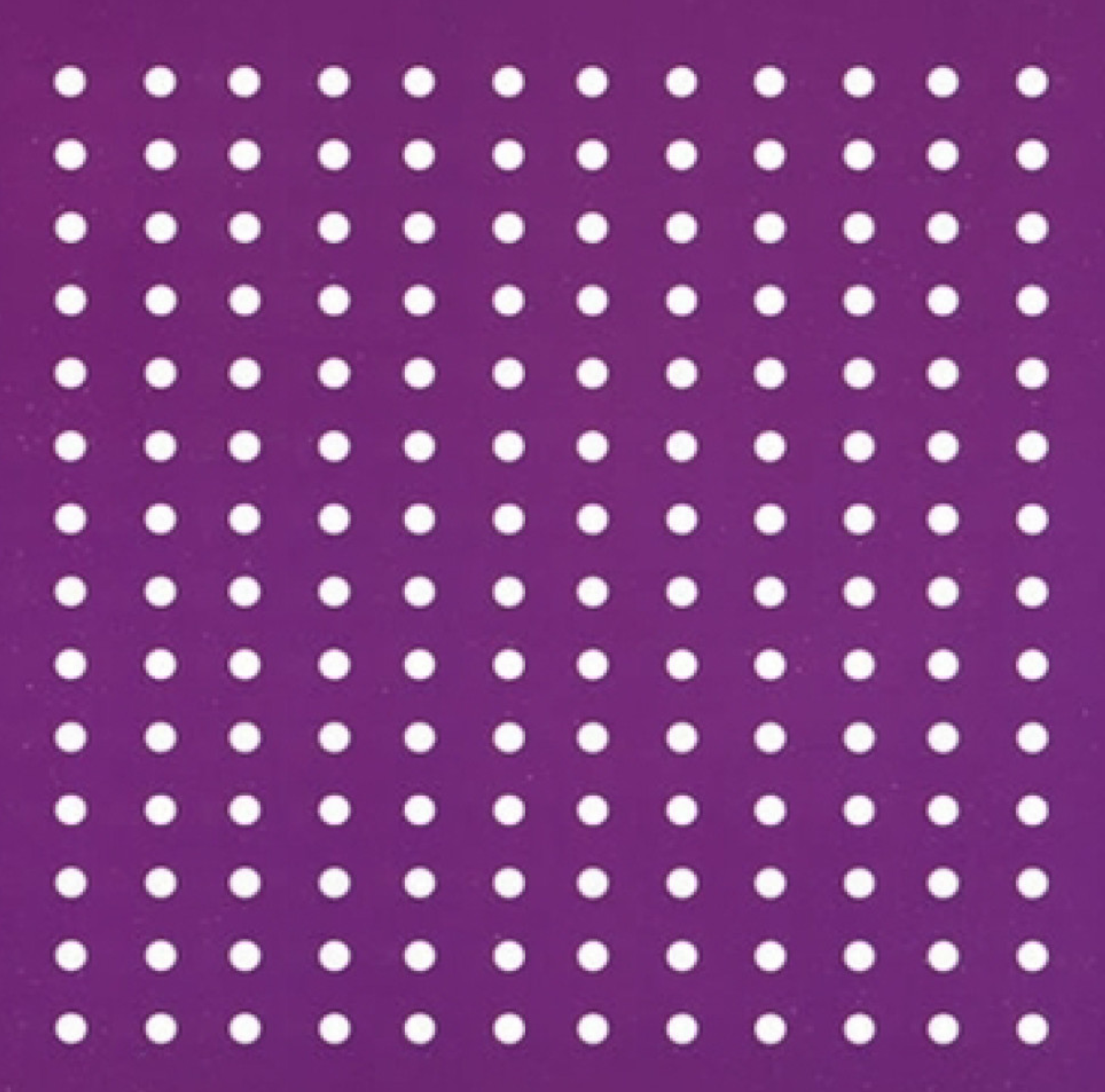
重点大学信息安全专业规划系列教材

无线自组织网络和对等网络
——原理与安全

易 平　吴 越　邹福泰　李建华　编著



清华大学出版社



无线自组织网络和对等网络原理与安全

易　平　吴　越　邹福泰　李建华　编著

清华大学出版社
北　京

内 容 简 介

本书详细阐述了无线自组织网络和对等网络的基本原理和安全技术。全书分为10章，内容包括无线自组织网络概述、无线自组织网络安全的研究进展、无线自组织网络安全架构、无线自组织网络中DoS攻击模型、无线自组织网络入侵检测研究、无线自组织网络的主动防护机制、无线局域网的安全、无线Mesh网络的安全、对等网络及研究进展、对等网络的安全问题。

本书适于作为通信与信息系统、电子与信息工程、计算机应用、计算机网络等相关专业的大学本科和研究生教材，也适合以上相关专业的应用开发人员、工程技术人员参考。

图书在版编目(CIP)数据

无线自组织网络和对等网络原理与安全/易平，吴越等编著. —北京：清华大学出版社，2009.8
ISBN 978-7-302-19933-5

Ⅰ. 无…　Ⅱ. ①易…　②吴…　Ⅲ. ①无线电通信—自组织系统—通信网—研究　②互联网络—研究　Ⅳ. TP92　TP393.4

中国版本图书馆CIP数据核字(2009)第056979号

责任编辑： 丁　岭　顾　冰
责任校对： 李建庄
责任印制：

出版发行： 清华大学出版社　　**地　　址：** 北京清华大学学研大厦A座
http://www.tup.com.cn　　**邮　　编：** 100084
社　总　机： 010-62770175　　**邮　　购：** 010-62786544
投稿与读者服务： 010-62776969，c-service@tup.tsinghua.edu.cn
质　量　反　馈： 010-62772015，zhiliang@tup.tsinghua.edu.cn
印 刷 者：
装 订 者：
经　　销： 全国新华书店
开　　本： 185×260　　**印　张：** 12.75　　**字　　数：** 306千字
版　　次： 2009年8月第1版　　**印　　次：** 2009年8月第1次印刷
印　　数： 1～　000
定　　价： .00元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题，请与清华大学出版社出版部联系调换。联系电话：010-62770177转3103　　产品编号：

FOREWORD

前 

进入20世纪90年代后，没有固定基础设施支撑、由若干移动节点组成的无线自组织网络，简称为移动Ad hoc网络(Mobile Ad hoc Networks)，逐渐成为分组无线网中的一个研究热点。无线自组织网络是一种不同于传统无线通信网络的技术。传统的无线蜂窝通信网络，需要固定的网络设备(如基地站)的支持，进行数据的转发和用户服务控制。而无线自组织网络不需要固定设备支持，各节点即用户终端自行组网，通信时，由其他用户节点进行数据的转发。这种网络形式突破了传统无线蜂窝网络的地理局限性，能够更加快速、便捷、高效地部署，适合于一些紧急场合的通信需要，如战场的单兵通信系统。它主要应用在抢险、抗灾、救援、探险、军事行动、应急任务和临时重大活动等，需要快速建立、移动、灵活的通信系统的场合中。它无论是在民用还是在军事上都有着显著的意义，而为了完成连续和无缝的通信要求，无线自组织网络将会起着至关重要的作用，因为仅仅基于现有的任何系统并不能支持更为广泛的、完全意义上的连续、无缝通信。在这一方面，无线自组织网络将是未来通信中关键而又现实的延伸，它可以灵活的扩展到任意的地域。

从2000年开始，对等网络(Peer-to-Peer Network，P2P网络)一直是计算机和互联网领域最受关注的热门话题之一。《财富》杂志将P2P技术列为影响Internet未来的四项科技之一，Intel公司也给了P2P技术极高的评价，将它称为“第三代网络革命”。对等网络将分布于世界各地的个人计算机组织起来，通过交换进行资源和服务的共享。这些资源和服务包括信息交换、高速缓存、处理能力、存储空间等。在对等网络中，每个节点自治又彼此依赖，自治是指每个节点独立决定自己的行为而不受其他例如集中式授权机构的控制，同时每个节点又需要相互协作获得信息资源、计算资源等。对等网络具有自组织特性，表现为网络具有高度的拓扑弹性和容错性。通常一个对等网络规模可达到几十万甚至上百万个计算节点，参与的计算节点以一种松散方式进行组织，非常适合广域网络和Internet的应用，并且已经涌现了不少非常具有影响力的应用系统，如PPLive、PPStream、BT、eMule、Skype等。事实上，对等网络已经在文件共享与内容分发、分布式数据存储、分布式计算、协同工作与服务共享、分布式深度搜索引擎、即时通信以及应用层多播等多个领域得到了广泛应用。对等网络不仅仅是一种技术体系，更是一种思想变革。它深刻影响了Internet的诸多应用，被视为下一代Internet应用的基础。

有感于无线自组织网络和对等网络技术的迅速发展,以及研究开发人员的对此领域进一步研究开发的需求,作者在自身研究工作积累的基础上精心编写了本书,让读者分享我们学习与研究工作的经验和成果。本书不仅可以使初学者能够了解领域内的研究现状,也可以为有一定研究基础的同行提供较为系统的相关的技术或方案,弥补这一领域内国内研究资料相对匮乏的境况,缩短国内学者的研究水平与国际一流水平的差距,贡献出更多的具有自主知识产权的研究成果。

本书共分12章。第1章介绍无线自组织网络的起源和发展。首先对无线自组织网络的概念和特点进行了简要叙述。然后介绍无线自组织网络的起源、发展历程和应用领域。其次着重阐述无线自组织网络领域中关键技术的研究现状及相关研究机构。

第2章介绍无线自组织网络的安全研究进展。由于无线自组织网络的独特结构,使得常规的安全方案无法应用,必须针对其特点设计专门的安全解决方案。本章从密钥管理、路由安全、入侵检测、增强合作几个方面介绍应用于无线自组织网络的安全解决方案。首先讨论密钥管理,主要介绍自组织的密钥管理和分布式的密钥管理两类算法,指出其优点和缺点。然后分析五种典型的路由安全协议,对它们进行综合比较并指出其存在问题及改进方法。接下来说明基于代理(agent)的分布式监视合作检测的入侵检测体系结构。最后讨论基于激励和基于惩罚的两种增强合作的机制。

第3章设计分析一种无线自组织网络的安全架构。安全架构借鉴免疫系统的思想,采用多代理来模拟实现淋巴细胞的免疫功能,设计一个适用于无线自组织网络的安全架构,实现对整个网络的入侵检测和入侵响应,同时还具有学习机制、分布式、自适应等特点。最后,通过实例分析阐明安全架构的有效性。

第4章讨论无线自组织网络中的攻击模型。本章分析无线自组织网络的安全弱点及其导致的各种DoS攻击方式。然后研究了按需路由协议AODV,发现其中的一些弱点,该弱点可能导致无线自组织网络中一种新的DoS攻击模型——Ad hoc Flooding攻击。该攻击主要针对移动Ad hoc网络中的按需路由协议,如AODV、DSR等。Ad hoc Flooding攻击是通过在网络中泛洪发送超量路由查询报文,大量地占用网络通信及节点资源,以至于阻塞节点正常的通信。分析Ad hoc Flooding攻击之后,提出邻居阻止的防御策略,即当入侵者发送大量路由查询报文时,邻居节点降低对其报文的处理优先级,直至不再接收其报文。

第5章介绍无线自组织网络的入侵检测。在无线自组织网络环境下,因为移动节点可能被攻击截获,从而泄露合法密钥,导致攻击从内部产生,传统的网络安全措施,如防火墙、加密、认证等技术,在无线自组织网络中难以应用。因此,只有通过入侵检测才能发现并清除入侵者。本章提出一种基于时间自动机分布式合作的入侵检测算法。首先,将整个网络分为各个监视区域,每一区域随机选出簇头担任监视节点,负责本区域的入侵检测。其次,按照DSR路由协议构筑节点正常行为和入侵行为的时间自动机,监视节点收集其邻居节点的行为信息,利用时间自动机分析节点的行为,发现入侵者。本算法不需要事先进行数据训练并能够实时检测入侵行为。

第6章分析无线自组织网络的主动防护机制。无线自组织网络是由移动节点自组织形成的网络,由于其动态拓扑、无线信道的特点,容易遭受各种安全威胁。至今提出的许多安

全方案主要集中于入侵阻止和入侵检测两个领域内。尽管这些安全方案能够取得一定的安全保障效果，但是它们都只是被动地去发现和阻止入侵者，并不能从根本上消除入侵行为。为了解决这个问题，本文提出一种自动入侵响应模型。该模型通过多种功能的代理组成一个整体来实现主动入侵响应，首先在每个节点布置监视代理，负责收集其周围每个邻居节点的行为信息。然后每个区域内的决策代理汇总监视代理的信息并进行判断。最后，阻击代理在入侵者周围形成一道移动防火墙，将入侵者包围并隔离于网络，消除入侵行为，从而实现无线自组织网络的主动防护机制。

第7章介绍无线局域网的安全。无线局域网(WLAN)是近年来发展迅速的无线数据通信网，但在发展同时，它又面临着许多安全问题。首先对无线局域网进行概述，然后对无线局域网的安全风险和安全需求进行分析，最后重点阐述无线局域网的安全技术以及安全协议。

第8章介绍无线Mesh网络安全，无线Mesh网络是一种多跳、具有自组织和自愈特点的网络，是近年来发展迅速的宽带无线通信网络，但它在发展的同时，又面临许多安全问题。本章首先对无线Mesh网络进行了概述，然后对无线Mesh网络的安全风险和安全需求进行了分析，最后重点阐述了基于MSA协议的安全协议及相关技术。

第9章介绍对等网络及其研究进展。对等网络系统是一个新兴的研究领域，近些年得到迅速发展。首先对对等网络进行概述性介绍，包括概念、分类及应用情况，然后重点介绍对等网络在路由、拓扑和查询这三方面的研究工作和研究进展。

第10章介绍对等网络的安全。对等网络由于其松散性的组织，随着应用的普及，其安全问题日益得到强调和重视。针对对等网络的节点的自私行为和恶意行为，介绍当前对等网络的激励机制、信任机制以及文件安全机制的研究工作。

本书系统、全面地介绍了无线自组织网络和对等网络中的一些新的模型、算法、协议、技术等，内容丰富、系统性强。在共同讨论形成本书大纲的基础上，易平撰写第1～6章，邹福泰撰写第7、9、10章，吴越完成第8章。全书最后由易平统稿。本书在编写过程中得到上海交通大学信息安全工程学院有关专家教授的关心与支持，在此向他们表示衷心的感谢。

作者衷心感谢清华大学出版社的大力支持，尤其感谢本书的编辑为本书付出的辛勤劳动。

无线自组织网络和对等网络涉及领域宽，内容多，发展快，本书的取材有些为学术界和工程技术界的研究成果，也包括本书作者的一些成果和观点。相关研究成果属于设计原作者，我们在书中均作了引用标识。我们尽量以客观的态度对待任何一项研究方法和成果，对于其中的争议甚至错误，希望读者去进一步甄别与探究。尽管我们力求完美，但作者水平有限，疏漏、不当与错误之处在所难免，欢迎读者批评指正。

本书得到国家自然科学基金(编号：60803060)、国家高技术研究发展863计划(编号：2006AA01Z436，2007AA01Z452，2009AA01Z118)、儿童发展与学习科学教育部重点实验室项目(编号：CDLS-2009-01)等项目的资助。

编　者

2009年6月于上海交通大学

CONTENTS

目录



第1章　无线自组织网络概述

CHAPTER 1

摘要：无线自组织网络技术是支持普适计算及未来移动通信系统的重要技术基础，对无线自组织网络相关技术的研究已经成为计算机网络和通信领域中的一个热点。本章首先对无线自组织网络的概念和特点进行了一个简要叙述。然后介绍了无线自组织网络的起源、发展历程和应用领域。其次着重阐述了无线自组织网络领域中关键技术的研究现状及相关研究机构。

关键字：无线自组织网络、概念、发展历程、应用领域、关键技术、研究机构。

1.1　研究背景

随着时间跨入21世纪，人类社会已进入一个崭新的发展阶段——信息社会。通信和网络技术的迅猛发展加速了信息交流，极大地促进了人类社会的“全球化”，深刻改变了社会的经济、政治与生活面貌。全球化的发展又进一步刺激了通信与网络技术的发展，人们追求任何人在任何时间、任何地点与任何人进行任何种类的信息交换。

在20世纪的大部分时间里，以固定电话网为代表的有线网络一直是信息的主要载体。然而在近二十年时间里，随着微电子技术与无线通信理论的迅速发展，无线通信网络获得了跨越式的发展，已成为全球通信网络的主要组成部分，最根本的原因在于无线通信网络使人们摆脱了通信线路的束缚，更接近于个人通信的需要。

近些年来，无线通信网络的发展非常迅速，这主要是由于个人通信的需求，无论是在支持范围上，还是种类、质量要求上都大大增加的缘故，而连接世界各地、可共享现有信息资源的因特网(Internet)的崛起更是极大地刺激了无线通信的发展。无线通信网络由于能快速、灵活、方便地支持用户的移动性而使它成为个人通信和Internet发展的方向，目前几乎所有的通信系统都与无线通信方式有关，如蜂窝系统、无绳系统、卫星通信系统、无线局域网与无线广域网(WLAN/WAN)[1]、移动IP[2]、无线ATM[3]、分组无线网(PRNET)[4]、无线自组织网络[5]等，而对无线和移动的相关研究成为这些通信系统中的最主要的部分。

传统意义上对无线通信网络的研究仅限于一跳无线网络，比如蜂窝系统

和无线局域网，它们都属于有基础设施的移动无线网络。在这些系统中，移动用户（或节点）在有限的区域里（即小区）移动，借助于固定的具有多部收发信机、可全双工方式工作的基站和可以大容量传输的有线骨干网络系统而与其他用户通信。当移动用户移出一个基站的覆盖范围而进入到另一个基站的覆盖范围内时由基站实现越区切换，这样移动用户就可以在整个通信网络中连续、无缝地通信。

进入 20 世纪 90 年代后，没有固定基础设施支撑、由若干移动节点组成的移动自组织网络，简称为无线自组织网络(Mobile Ad hoc Networks)，逐渐成为分组无线网中的一个研究热点。无线自组织网络独立于任何静态的基础设施，可即时建立。它主要应用在抢险、抗灾、救援、探险、军事行动、应急任务和临时重大活动等，需要快速建立、移动、灵活的通信系统的场合中。它无论是在民用还是在军事上都有着显著的意义，而为了完成连续和无缝的通信要求，无线自组织网络将会起着至关重要的作用，因为仅仅基于现有的任何系统并不能支持更为广泛的、完全意义上的连续、无缝通信。在这一方面，无线自组织网络将是未来通信中关键而又现实的延伸，它可以灵活的扩展到任意的地域。

无线自组织网络是一个复杂系统，所涉及的研究内容非常广泛，目前对它的研究和应用已发展成为通信领域的一个独立分支，存在一些需要彻底研究的问题。

1.1.1 无线自组织网络的概念及特点

无线自组织网络是由具有无线通信能力移动节点组成的、具有任意和临时性网络拓扑的动态自组织网络系统，其中每个节点既可作为主机也可作为路由器使用。Ad hoc 的意思是 for this 引申为 for this purpose only，即“为某种目的设置的，特别的”意思，即 Ad hoc 网络是一种有特殊用途的网络。移动终端具有路由功能，可以通过无线连接构成任意的网络拓扑，这种网络可以独立工作，也可以与 Internet 或蜂窝无线网络连接。在后一种情况中，无线自组织网络通常是以末端子网的形式接入现有网络。考虑到带宽和功率的限制，无线自组织网络一般不适于作为中间传输网络，它只允许产生于或目的地是网络内部节点的信息进出，而不让其他信息穿越本网络，从而大大减少了与现存 Internet 互操作的路由开销。无线自组织网络中，每个移动终端兼备路由器和主机两种功能：作为主机，终端需要运行面向用户的应用程序；作为路由器，终端需要运行相应的路由协议，根据路由策略和路由表参与分组转发和路由维护工作。在无线自组织网络中，节点间的路由通常由多个网段（跳）组成，由于终端的无线传输范围有限，两个无法直接通信的终端节点往往要通过多个中间节点的转发来实现通信。所以，它又被称为多跳无线网、自组织网络、无固定设施的网络或对等网络。无线自组织网络同时具备移动通信和计算机网络的特点，可以看作是一种特殊类型的移动计算机通信网络。

图 1-1 描述了一个由五个主机组成的简单的无线自组织网络。主机 D 不在主机 A 的无线覆盖范围之内（用环绕主机 A 的圆环表示），同时主机 A 也不在主机 D 的无线覆盖范围内。如果主机 A 和 D 之间需要交换信息，就需要主机 B、C 为它们转发分组，因为主机 B、C 在主机 A 和 D 的无线覆盖范围之内。

与通常的网络相比，无线自组织网络具有以下特点[5]：

(1) 网络的自组织性：无线自组织网络相对常规通信网络而言，最大的区别就是可以在任何时刻、任何地点不需要硬件基础网络设施的支持，快速构建起一个移动通信网络。它

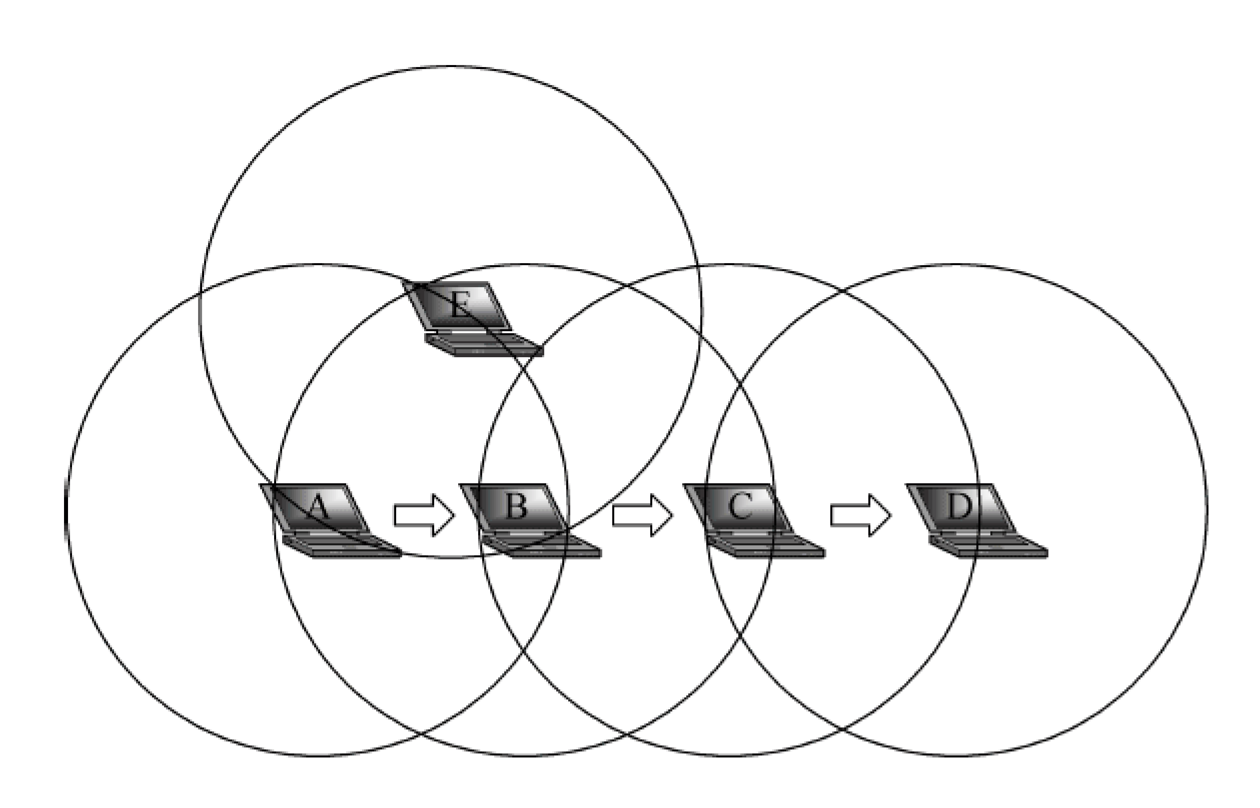


图 1-1　一个简单的无线自组织网络

的建立不依赖于现有的固定网络通信设施，由网络本身节点自组织形成网络。无线自组织网络的这种特点很适合灾难救助、偏远地区通信等应用。

(2) 动态的网络拓扑结构：在无线自组织网络中，移动主机可以在网中以任意速度和任意方式移动，主机的移动会导致主机之间的链路增加或消失，主机之间的关系不断发生变化。加上无线发送装置发送功率的变化、无线信道间的互相干扰及地形和地物等综合因素影响，各移动节点间的连接关系将时刻发生变化，因此，造成网络拓扑结构不断发生变化，而且变化的方式和速度都是不可预测的。对于常规网络而言，网络拓扑结构则相对较为稳定。

(3) 多跳的通信路由：由于节点无线发射功率的限制，节点的覆盖范围有限。当它要与其覆盖范围之外的节点进行通信时，需要中间节点进行转发。此外，无线自组织网络中的多跳路由是由普通节点协作完成的，而不是由专用的路由设备(如路由器)完成的。网络中每一个节点可充当多个角色，它们可以是服务器、终端、路由器。

(4) 有限的无线通信带宽：在无线自组织网络中没有固定基础设施的支持，因此，主机之间的通信均通过无线传输来完成。由于无线信道本身的物理特性，它提供的网络带宽相对有线信道要低得多。除此以外，考虑到竞争共享无线信道产生的碰撞、信号衰减、噪音干扰等多种因素，移动终端可得到的实际带宽远远小于理论中的最大带宽值。

(5) 有限的主机能源：在无线自组织网络中，主机均是一些移动设备，如 PDA、便携计算机或掌上电脑。由于主机可能处在不停的移动状态下，主机的能源主要由电池提供，因此，网络具有能源有限的特点。

(6) 网络的分布式特性：在无线自组织网络的各节点都具备独立的路由能力，没有中心控制节点对各节点网络操作进行控制，节点通过分布式协议互联。一旦网络的某个或某些节点发生故障，其余的节点仍然能够正常工作。

(7) 生存周期短：无线自组织网络主要用于临时的通信需求，相对与有线网络，它的生存时间一般比较短。

(8) 安全性较差：无线自组织网络是一种特殊的无线移动网络，由于采用无线信道、有限电源、分布式控制等技术，它更加容易受到被动窃听、主动入侵、拒绝服务、剥夺“睡眠”等网络攻击。信道加密、抗干扰、用户认证和其他安全措施都需要特别考虑。

(9) 移动节点的局限性：无线自组织网络中，移动节点具有携带方便、轻便灵巧等好

处,但是也存在固有缺陷,例如能源受限、内存较小、CPU 性能较低等,从而给应用程序设计开发带来一定的难度,同时屏幕等外设较小,不利于开展功能较复杂的业务。

1.1.2 无线自组织网络的发展历程

无线自组织网络技术起源于 20 世纪 70 年代,它是在美国国防部高级研究计划局(DARPA)资助研究的“战地无线分组数据网(PRNET)”[4]项目中产生的一种新型网络技术。DARPA 当时所提出的是一种军用无线分组数据通信网络。在此之后,DARPA 于 1983 年启动了高残存性自适应网络项目 SURAN(Survivable Adaptive Network)[6],研究如何将 PRNET 的研究成果加以扩展,以支持更大规模的网络。1994 年,DARPA 又启动了全球移动信息系统 GloMo(Globle Mobile Information Systems)项目[7],旨在对能够满足军事应用需要的、可快速展开、高抗毁性的移动信息系统进行全面深入的研究,以便能够建立某些特殊环境或紧急情况下的无线通信网络。无线自组织技术就是吸取了 PRNET、SURAN 以及 GloMo 等项目的组网思想,而产生的一种新型的网络结构技术。美国军方一直在研究适用于军方的无线自组织网络技术,后来又陆续资助了联合战术无线系统(JTRS)[8]等项目。成立于 1991 年 5 月的 IEEE 802.11 标准委员会[9]采用了“Ad hoc 网络”一词来描述这种特殊的自组织对等式多跳移动通信网络,无线自组织网络就此诞生。IETF 也将无线自组织网络称为 MANET(Mobile Ad hoc Networks)[10]。

随着移动通信和移动终端技术的高速发展,无线自组织网络技术不仅在军事领域中得到了充分的发展,而且也在民用移动通信中得到了应用。典型的系统有加拿大最早研究的业余分组无线网(TAPR)[11],图书馆自动化无线电网络[12]等。Internet 任务工作组(IETF)于 1996 年成立了 MANET(Mobile Ad hoc Networks)工作组,专门研究 Ad hoc 网络环境下基于 IP 协议的路由协议规范和接口设计[10]。这使得无线自组织网络的设计思路也由传统的单一技术体系过渡到基于 IP 的多技术体系,从而使该网络更具有开放型、适应性、灵活性,提高了开发和应用速度。随着配备有无线收发设备的高性能移动终端的降价和将要随之而来的普及性,加上人们对于个人通信需求的日益增长,使得无线自组织网络的研究重新开始得到国内外研究人员的重视。特别是从 1998 年以来,无论是国内还是国外,各科研团体对无线 Ad hoc 网络的研究不断升温,尤其是在网络层的路由协议方面,其研究工作已经取得了很大的进展。

1.1.3 无线自组织网络的应用领域

无线自组织网络的许多优良特性为它在民用和军事通信领域占据一席之地提供了有利的依据。首先,网络的自组织性提供了廉价而且快速部署网络的可能。其次,多跳和中间节点的转发特性可以在不降低网络覆盖范围的条件下减少每个终端的发射范围,从而降低设计天线和相关发射/接收部件的难度,也降低了设备的功耗,从而为移动终端的小型化、低功耗提供了可能。从共享无线信道的角度来看,无线自组织网络降低了信号冲突的几率,提高了信道利用率。从对使用者的保护来看,高功率的无线电波产生的电磁辐射对用户的身体健康也有影响。另外,网络的鲁棒性、抗毁性满足了某些特定应用需求。它的应用场合可以归纳为以下几类。

(1) 军事应用：军事应用是无线自组织网络技术的主要应用领域。在现代化的战场上，由于没有基站等基础设施，装备了移动通信装置的军事人员、军事车辆以及各种军事设备之间可以借助无线自组织网络进行信息交换，以保持密切联系、协作完成作战任务。装备音频传感器和摄像头的军事车辆和设备也能够组成无线自组织网络将在目标区域收集重要的位置和环境信息传送到处理节点。另外，需要通信的舰队战斗群之间也可以通过无线自组织网络建立通信而不必依赖陆地或卫星通信系统。无线自组织网络因其特有的无需架设网络设施、可快速展开、抗毁性强等特点，它是数字化战场通信的首选技术，并已经成为战术互联网的核心技术。为了满足信息战和数字化战场的需要，美军研制了大量的无线自组织网络设备，用于单兵、车载、指挥所等不同的场合，并大量装备部队。美军的近期数字电台NTDR和无线互联网控制器等通信装备都使用了无线自组织网络技术。

(2) 移动会议：目前，越来越多的人携带手提电脑、PDA等便携式设备参加各种会议。如果与会者不用借助路由器、集线器或基站就能将各种移动终端快速地组织成无线网络从而完成提问、交流以及资料的分发，这无疑具有重要的意义，而无线自组织网络就具有这样的功能。当一些移动用户聚集在办公室外的某个环境时，他们也可以借助无线自组织网络来协同工作。此外，借助无线自组织网络还可以实现分布式会议。

(3) 紧急和突发场合：在自然灾害或其他各种原因导致网络基础设施出现故障或无法使用时，快速地恢复通信是非常重要的。借助于无线自组织网络技术和协议，可以快速地建立临时网络，延伸网络基础设施，从而为营救赢得时间，减少灾难所带来的危害。例如，在因发生了地震、水灾、火灾或遭受其他灾难后而使得基站、通信干线等基础通信设施无法使用时，可以形成无线自组织网络来快速地建立联系，组织营救。此外，当刑警或消防队员紧急执行任务时，可以通过无线自组织网络来保障通信指挥的顺利进行。

(4) 偏远野外地区：当处于边远或野外地区时，无法依赖固定或预设的网络设施进行通信。无线自组织网络技术具有单独组网能力和自组织特点，是这些场合通信的最佳选择。

(5) 临时场合：无线自组织网络的快速、简单组网能力使得它可以用于临时场合的通信。比如庆典、展览等场合，可以免去布线和部署网络设备的工作。

(6) 动态场合和分布式系统：通过无线连接远端的设备、传感节点和激励器，无线自组织网络可以方便地用于分布式控制，特别适合于调度和协调远端设备的工作，减少分布式控制系统的维护和重配置成本。无线自组织无线网络还可以用于在自动高速公路系统(AHS)中协调和控制车辆[12]，对工业处理过程进行远程控制等。

(7) 个人通信：个人局域网(PAN)是无线自组织网络技术的又一应用领域，用于实现PDA、手机、掌上电脑等个人电子通信设备之间的通信，并可以构建虚拟教室和讨论组等崭新的移动对等应用(MP2P)。考虑到电磁波的辐射问题，个人局域网通信设备的无线发射功率应尽量小，这样无线自组织网络的多跳通信能力将再次展现它的独特优势。

(8) 商业应用：组建家庭无线网络、无线数据网络、移动医疗监护系统和无线设备网络，开展移动和可携带计算以及无所不在的通信业务等。

(9) 其他应用：考虑到无线自组织网络具有很多优良特性，它的应用领域还有很多，这需要我们进一步去挖掘。比如它可以用来扩展现有蜂窝移动通信系统的覆盖范围[13]，实现地铁和隧道等场合的无线覆盖，实现汽车和飞机等交通工具之间的通信，用于辅助教学和构建未来的移动无线城域网和自组织广域网[14]等。

1.2 无线自组织网络的主要研究领域

无线自组织网络拥有许多优良特性,能够广泛应用于许多场合,但同时还应注意到设计和应用该网络面临的诸多技术难点。随时变化的链路特性和网络拓扑、节点的移动性、受限的链路带宽和节点能量、恶劣的无线环境和安全性都是我们必须面对的问题。因此,必须仔细考虑节点的硬件设计(小型化、智能化和节能化)和协议栈的各个层次。在物理层,要解决衰落、多径干扰、功率控制等无线通信经常遇到的问题。在数据链路层,要解决多跳共享的广播信道的有效接入问题。网络层需要特殊的路由协议来维护网络动态变化的拓扑信息。在传输层,要解决无线环境下传输层的效率问题。应用层要具有一定的自适应流量控制功能。并且还要考虑协议栈各个层次的紧密协作,以适应网络条件和应用需求的变化。此外,网络的自组织特性、无中心控制、易配置性和可编程性等特征都对协议的设计提出了新的特殊的要求。也就是说,无线自组织网络的设计需要综合考虑资源效率、能量保护、信道接入、路由和资源分配等问题并进行合理的折中。

无线自组织网络一方面作为自治系统,有自身特殊的路由协议和网络管理机制;另一方面作为 Internet 在无线和移动范畴的扩展和延伸,它又必须能够提供到 Internet 的无缝的接入机制。当前 Internet 已经可以在一定程度上保证综合业务传输的服务质量,近年来随着多媒体应用的普及和 Ad hoc 网络在商业应用的进展,人们很自然地会产生在无线自组织网络上传送综合业务的需求,并且希望能像固定的有线网络一样为不同业务的服务质量提供保障,因此无线自组织网络对 QoS 保障的支持显得越来越迫切和重要。但是与固定的有线网络不同,在无线自组织网络中提供 QoS 支持将面临许多不同于传统网络的新问题和挑战。

无线自组织网络是一种动态变化的基于无线信道的自组织网络,它的体系结构、QoS 保障和应用等问题比较复杂并难以实现。传统固定网络和蜂窝移动通信网中使用的各种协议和技术无法被直接使用,因此需要为无线自组织网络设计专门的协议和技术。目前,无线自组织网络的主要研究领域包含 MAC 层协议、路由协议、如何节省能源、QoS 保证、网络安全、多播、网络管理等多个方面。

1.2.1 MAC 层协议

在无线自组织网络中,由于节点的通信范围有限和随机移动特性,将会产生隐藏和暴露终端等问题[15,16]。"隐藏终端"是指,在图 1 中,当节点 A 向节点 B 发送信息时,节点 C 未侦测到 A 也向 B 发送,故 A 和 C 同时将信号发送至 B,引起信号冲突,最终导致发送至 B 的信号都丢失了。此时称 A 对 C 是隐藏的终端。"暴露终端"是指,正当 B 向 A 发送信息时,C 要向 D 发送信息,但此时 C 监听到 B 正在发送报文,C 认为信道忙,于是不能向 D 发送,此时实际上 C 是可以向 D 发送信息的。这就是暴露终端问题。

由于现有的 CSMA 协议不能直接应用于无线自组织网络,因此必须采用新型的 MAC 协议,以便获得较高的信道利用率和较低时延的公平接入。对于隐藏终端问题,可以采用控制报文握手协议的方法来解决。但是在单信道条件下则无法彻底解决,为此必须采用双信道方式:收发数据的数据信道和收发控制信号的控制信道。目前,已经提出了一些可以应

用于无线自组织网络的 MAC 协议[17]，如 MACA[18]、MACAW[19]、FAMA[20] 和 DCF，以及 DBTWA[21] 等。其中，多址接入冲突避免协议（MACA）采用了 RTS/CTS 信道的握手机制，提高了无线信道的利用率，并解决了部分隐藏终端问题，但是仍然无法避免控制分组间的冲突，不具备链路层的确认机制，协议的公平性也较差。针对 MACA 存在的问题，无线多址接入冲突避免协议（MACAW）进行了改进，采用乘法增加线性减少退避算法替代了二进制指数退避算法，可以获得更好的公平性，同时采用 RTS 和 CTS 的数据应答握手机制，进一步提高信道的利用率，但是协议的主要缺点是通信中控制信息交互次数太多，并且也不能完全解决暴露终端问题。FAMA 是基于单信道的无线自组织网络信道接入协议中较成功的一种。美军在无线互联网网关中使用的信道接入协议就是 FAMA，它对 MACA 和 MACAW 作了进一步改进，通过延长 RTS 和 CTS 控制报文的长度来消除控制报文的冲突，从而比较好地解决了隐藏终端问题，同时节点可发送多个报文，增加了网络的吞吐量。IEEE 802.11 的 DCF 提供了对无线自组织网络的支持，信道接入采用 CSMA/CA 机制，并采用了类似于 MACAW 的握手机制，但区别在于它使用了载波监听功能。以上的几种方法都是建立在所有相关节点都能听到 RTS/CTS 消息的假设条件基础上，然而在高速移动的大型网络中这种假设有时并不成立。当网络负载很高时 CTS 发生冲突的概率很大，为了解决这些问题，提出了双忙音多址接入协议 DBTMA。通过双信道加忙音的方法，不仅解决了隐藏终端问题，而且降低了控制信号发生冲突的概率，因而网络利用率较 MACAW 提高一倍。

上述的信道接入协议，在一定程度上解决了无线自组织网络中的信道接入问题，但都存在着一定的局限性，一般都要求较小的网络规模和较低的移动性。其中，DBTMA，是相对比较好的信道接入控制方案，有着良好的应用前景。当前，信道接入协议的研究仍在进行之中，未来的研究方向包括如何为实时业务提供较好的支持、如何支持广播和多播业务以及考虑支持业务优先级和基于受控方式的接入机制等。

1.2.2　路由协议

无线自组织网络路由协议方面的研究是无线自组织网络最重要的研究方向之一。网络中节点随时移动、网络拓扑动态变化，使得传统的距离向量和链路状态路由协议不再适用于无线自组织网络，必须根据无线自组织网络的特点进行修改，或者提出一些新的路由协议。设计良好的路由协议是建立无线自组织网络的首要问题，也是主要的研究热点和难点。至今已经提出了数十种适用于无线自组织网络的路由协议[22~25]，综合起来可以分为三大类：平面路由协议、层次路由协议、位置路由协议。在平面路由协议中，各节点具有平等的地位。而在层次路由协议中，通常指定一些节点担任簇头的角色，负责其区域内节点和区域间节点的通信。位置路由则要求每个节点必须装备 GPS 设备，能够确定节点方位。平面路由协议中有可分为预先路由协议和按需路由协议。每类协议中又包含许多协议，如图 1-1 所示。由于篇幅所限，这里只提出了一些比较典型的路由协议（见图 1-2）。如：预先路由协议中的 FSR[26]、DSDV[27]、OLSR[28]、TBRPF[29]。按需路由协议中的 DSR[30]、AODV[31]。层次路由协议中的 HSR[32]、CGSR[33]、LANMAR[34]、ZRP[35]。位置路由协议中的 GeoCast[36]、DREAM[37]、LAR[38]、GPSR[39]。能量路由协议中的 PARO[40]、PAMAS[41]。

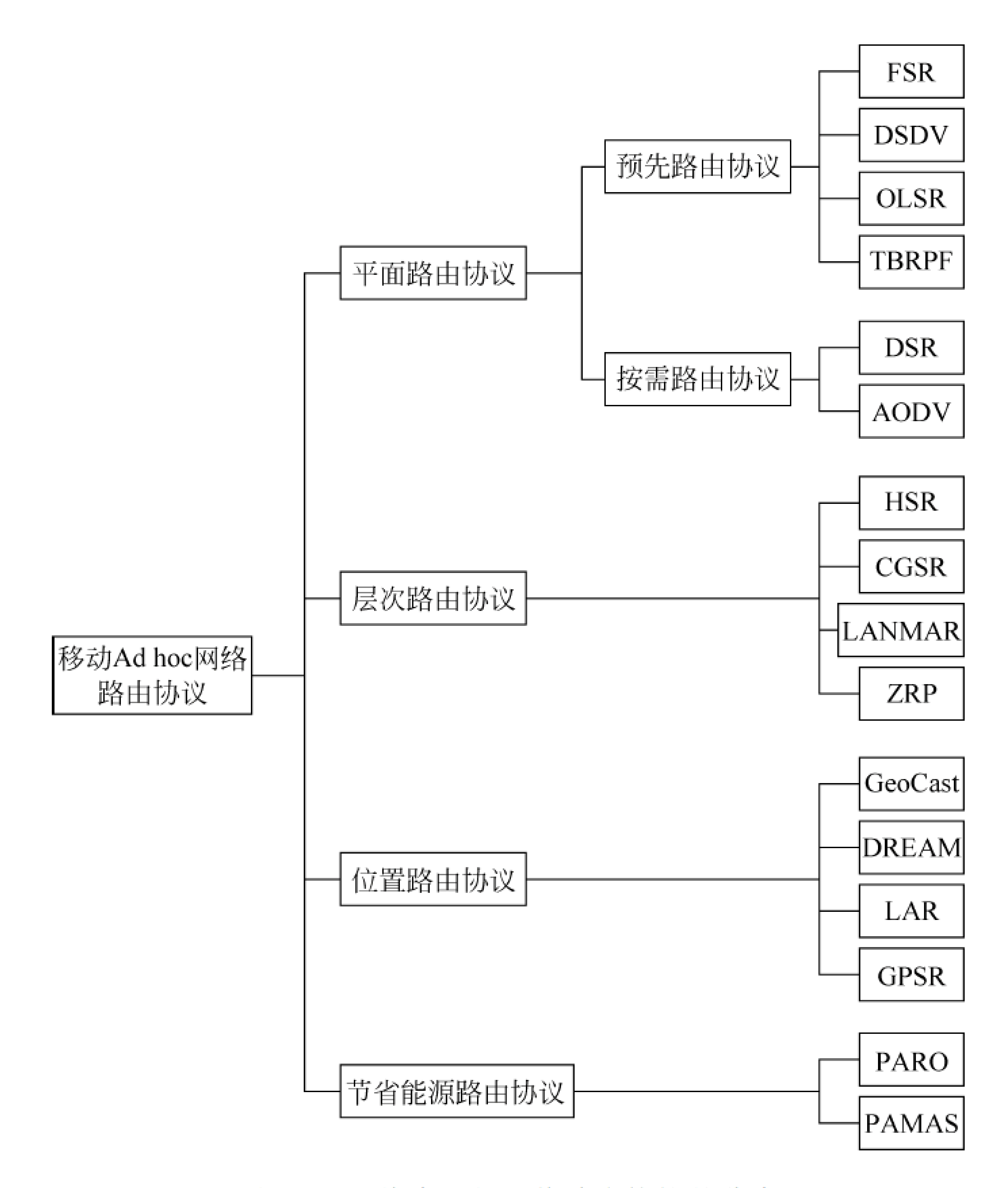


图 1-2　无线自组织网络路由协议的分类

平面路由协议的网络比较简单，无需任何的结构维护过程。源节点和目的节点之间可以存在多条路径，可以使用多条路径实现负荷分担，也可以为不同的业务类型选择适当的路径。网络中所有节点是对等的，原则上不存在瓶颈，所以比较健壮。平面路由协议中节点的覆盖范围比较小，相对较安全。平面路由协议的最大缺点是网络规模受限。在平面路由中，每一个节点都需要知道到达其他所有节点的路由。由于节点的移动性，维护这些动态变化的路由信息需要大量的控制消息。网络规模越大，路由维护的开销就越大。当网络的规模增加到某个程度时，所有的带宽都可能会被路由协议消耗掉。所以平面路由协议网络的可扩充性较差。

平面路由协议中可分为预先路由协议和按需路由两类，预先式路由协议一般是表驱动的，它需要在每个节点维护一个或多个路由表，其中包含了该节点到网络中所有其他节点一致的、最新的路由信息。为了维护这样的路由表，每个节点要定期向网络广播拓扑信息，以维护一致的网络视图。采用不同数量和内容的路由表以及不同的广播策略，即形成了各种不同的具体路由协议，如：DSDV[27]、FSR[26]、OLSR[28]、TBRPF[29]。DSDV[27]原理是基于经典的 Bellman-Ford 路由机制，其所做的主要改进是防止路由表产生循环路由。在 DSDV (Destination Sequenced Distance Vector Routing)协议中，每个节点维护一个路由表，其中记录了网络中所有其他节点以及到达这些节点的跳数。路由表中的记录由目的节点指定的

顺序号标识,该顺序号隐含了时间顺序信息,以区分新路由和过时路由,并由此避免路由循环。FSR(Fisheye State Routing)[26]是预先型链路状态路由协议,其目的是通过鱼眼效应,即近处的物体清晰,远处的物体模糊,来减少路由信息流量。FSR 对链路状态算法进行了一些修改。仅在邻节点间交换链路状态信息,而不是将链路状态信息广播到整个网络。FSR 对于路由表中不同的记录采用不同的时间间隔交换链路状态信息。对于较近的节点用较短间隔交换链路状态信息,对于较远的节点用较长的间隔交换状态信息。通过这些手段减少了路由信息,并降低了传输频率。OLSR(Optimized Link State Routing)[28]是一种优化的链路状态协议,与其他表驱动的预选型路由协议一样,节点间需要有规律地交换网络拓扑信息。但不是每个节点都可以交换路由信息。被邻节点选为多点中继站的节点才能周期性地向网络广播控制信息,控制信息中包含了把它选为中继节点的那些节点的信息。以告诉网络中其他节点与这些节点直接相连。只有中继节点被用做路由节点。非中继节点不参与路由计算,不转发路由信息。OLSR 减少了路由信息在网络中泛洪程度,降低了网络负载。

按需路由协议的出发点是只有当节点需要路由时才建立路由,通信过程中才维持路由,通信完毕就不再维持路由。一般地,按需路由都包括三个过程:路由发现过程、路由维持过程和路由拆除过程。DSR(Dynamic Source Routing)[30]使用了源路由机制,每一个分组的分组头中包含整条路由的信息,其优点是中间节点不需要维持当前的路由信息,分组自己带有路由信息。再加上按需路由的特性,就避免了周期性的路由广播和邻节点的检测。AODV(Ad hoc On Demand Distance Vector Routing)[31]是建立在 DSDV 算法之上的,但是它并不维持一个路由表,而是在需要的时候才启动路由选择过程,因此大大地降低了路由维持的开销。事实上它是 DSR 和 DSDV 的组合,它借用了 DSR 的按需路由发现和路由维持机制,利用了 DSDV 的多跳路由、顺序编号和周期更新的机制。

预先路由协议通过连续地检测链路质量,时刻维护准确的网络拓扑和路由信息。优点是发送报文时可以立即得到正确的路由。但预先路由协议需要大量的控制报文,开销太大。按需路由协议中的节点不用持续维护网络的拓扑结构。仅当需要时,才查找相应的路由,这就节省了路由维护的开销,特别是当网络负荷不是很重时,节省的开销更加可观。但查找路由会引入较大的时延,不适用于时延敏感型应用。

层次路由协议中,簇内成员的功能比较简单,基本上不需要维护路由,这大大减少了网络中路由控制信息的数量。簇头节点功能要复杂一些,它要维护好到达其他簇头的路由,还要知道所有节点与簇的所属关系。但总的来说,在相同网络规模的条件下路由开销要比平面结构的小。如果簇内通信的信息量占较大比例时,各簇可以互不干扰地进行通信,系统的吞吐量显然要比平面结构的要高。层次路由协议的最大优点是可扩充性好,网络规模不受限制。必要时可以通过增加簇的个数或级数来提高网络的容量。但是层次路由也有它的缺点。首先,维护层次路由需要较复杂的簇头选择算法,簇头选择算法需要仔细设计。其次,簇间的信息都要经过簇头寻路,不一定能使用最佳路由。比如在不同簇中但互为邻居的节点,在平面结构中可以直接通信,但分簇后要通过两个簇的簇头转交。HSR(Hierachical State Routing)[32]是一种基于簇头的多层链路状态路由协议。它不仅将网络在物理上划分为一个个区域,而且将网络逻辑上分成多个层次,低一层的簇头是高一层的成员。分层的目标是减少路由维护的信息量。CGSR(Clusterhead-Gateway Switch Routing)[33]是一种基于

簇头的两层距离向量路由协议。整个网络节点划分为许多区域，每个区域选出一个簇头，簇头负责与簇内节点之间通信。两个区域交界的节点为网关节点，其负责跨区域报文转交。整个报文的传送协过程为：源节点—簇头—网关节点—簇头—网关节点…簇头—目标节点。ZRP(Zone Routing Protocol)[35]是一个分层路由协议，它巧妙地结合了按需路由协议和预先路由协议的长处。ZRP将整个网络分成若干个区域，在一个区域内的路由操作采用预先路由方案，用于维护区域内节点可达性的完整信息。在区域之间使用按需路由。

全球定位系统(Global Positioning System，GPS)技术的迅速发展，使得定位精度可达到数米的误差范围之内。一些研究表明，在地理位置信息的帮助下，无线自组织网络路由协议的性能可以得到明显提高。下面介绍几种基于位置信息的路由协议。GeoCast (Geographic Addressing and Routing)[36]协议中，通过分层的位置路由器来实现其报文的转发。当源节点发送报文时，首先看其目标节点是否在本区域内，如果不在则交上层位置路由器，以此类推，直到转发到其节点所在区域位置路由器，再由位置路由器发给区域内所有节点，包括目标节点。DREAM(Distance Routing Effect Algorithm for Mobility)[37]是一个预先路由协议，使用位置信息确定数据分组的泛洪方向，它通过给控制信息设置不同的TTL(Time To Live)值来实现所谓的距离效应，即两个节点相距越远它们的相对运动似乎越慢，从而减少网络中的控制信息。LAR(Location Aided Routing)[38]是一个类似于DSR的随选型路由协议，利用位置信息选择控制分组泛洪的方向，提出了两种方案——创建包含源节点的请求区域以及利用分组在传输过程中离目的节点越来越近的特点。

无线自组织网络与传统网络的重要差别之一是其节点都具有移动能力，这要求设备尽量小型化，而且需要节省能量的使用，以延长节点寿命，进而延长整个网络的寿命。路由协议的合理设计可以降低能量的消耗。PARO[40]协议利用增加多点中继转发来降低传送报文的能量消耗。长距离的发送报文会消耗较多的能量。当两距离较远的节点A、B进行通信时，尽管它们都在相互直接通信范围之内，PARO会在节点A、B之间寻找一些中继节点进行转发，缩短单跳的传输距离，从而减少整个过程的能量消耗。PAMAS(Power Aware Multi-Access Protocol)[41]提出了一种减少能量消耗的信道访问协议。它利用了下述无线通信的特点，当源节点与目标节点进行报文传输过程中，周围在其通信影响范围内的节点都只能被动收听，不能发送。此时，如果无关节点将其电源关掉，将会节省许多能源。按照上述原理，PAMAS提出，当节点没有报文发送时，而信道又忙时，它应关掉其电源到报文发送结束。当节点有报文发送时，而信道又被占用时，它也应关掉其电源一段时间用于退避。

本节较全面地回顾了无线自组织网络相关的各种路由协议，根据不同的策略对协议进行了分类，分为平面路由协议、层次路由协议、位置路由协议和节点能源路由协议。并对各种类型的协议进行了分析和比较，指出了各自的特点。目前的各种路由算法各有其长处和短处，各有其适用的场合，没有一种算法能适合于所有情况。

1.2.3 多播路由协议

在无线自组织网络的应用中，如灾难救助、战场指挥、临时会议，通常都有一个共同的需求，就是一到多或是多到多的数据传输，因此，多播路由协议在无线自组织网络中具有非常重要的作用。近几年，研究人员提出了一些能适应Ad hoc环境的多播路由协议，这些协议

可以大致分成三类[42]：一类是基于树(tree-based)的，如 AMRIS[43]、MAODV[44]、LAM[45]、LGT[46]，它们在源和接收者之间只提供一条路由；第二类是基于格网(mesh-based)的，如 ODMRP[47]、CAMP[48]、FGMP[49]，它们在传输数据时能在源和接收者之间提供多条路由；第三类是混合多播，如 AMRoute[50]、MCEDAR[51]，如图 1-3 所示。

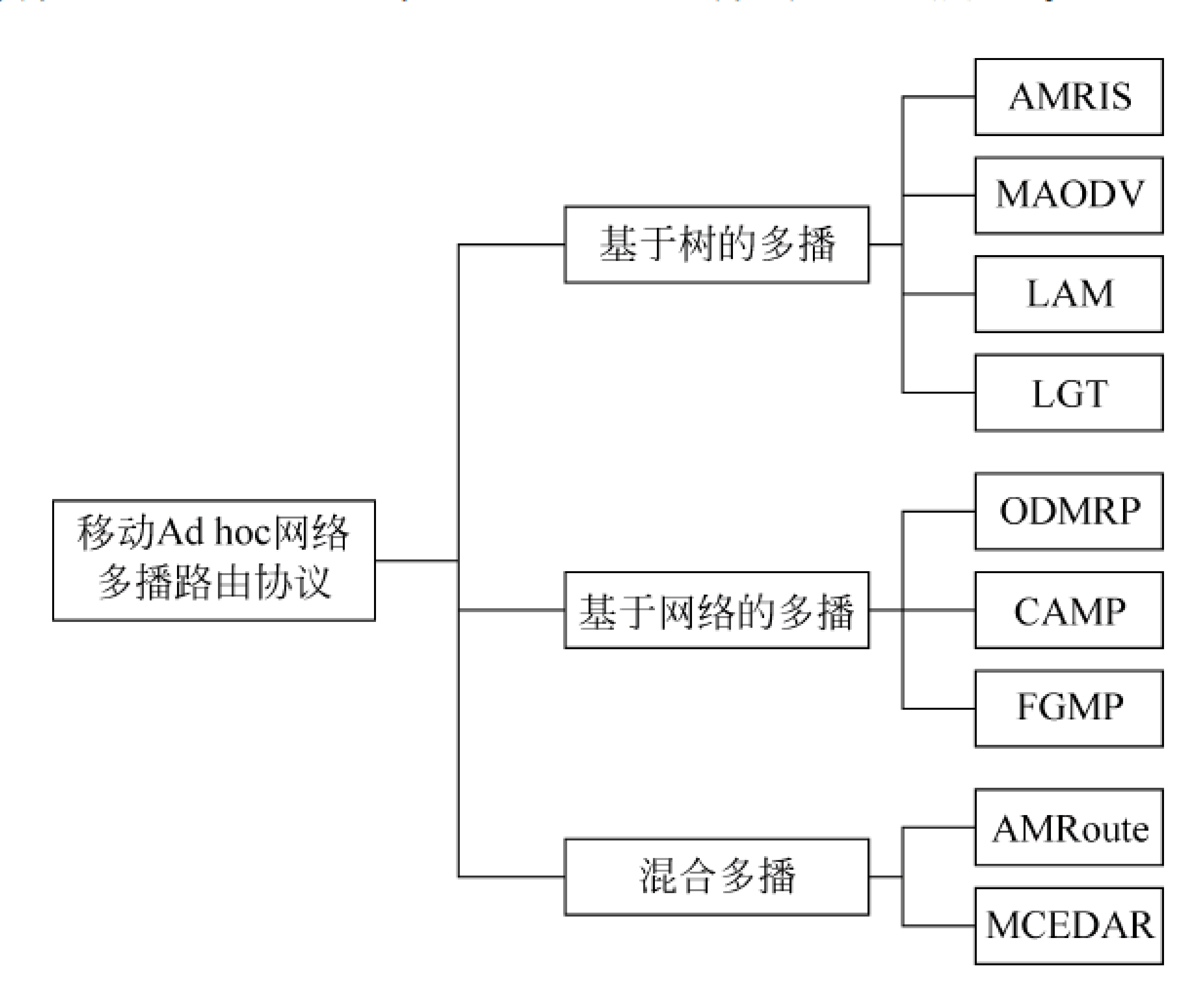


图 1-3　多播路由协议分类

在 Ad hoc 环境中，基于树的多播协议来源于有线网络多播协议，它们通过建立多播共享树的方法来实现多播。由于网络拓扑结构经常会发生变化，因此基于树的多播路由协议数据递交率一般都较低，不能满足应用的需要，但状态维护的工作量较少。而基于网格的路由协议因为能提供冗余路径，有较高的健壮性。它们较为适用于动态的拓扑结构，分组递交率一般都较高，但状态维护的工作量较多。结合两种方法，形成第三种多播方式——混合多播，核心部分采用网格方式，外围节点采用基于树的方式。

在这些多播路由协议中，ODMRP 由于是通过源节点在全网广播控制信息来建立和维护路由的。因而，当源节点个数较多时，协议将产生较大的开销，这一特性大大影响了协议的可扩展性。为了减小这种广播所带来的控制开销，CAMP 使用了核心网格，并不在全网泛洪，模拟结果表明它的可扩展性好于 ODRMP，不过，该协议依赖于底层的单播路由协议。

AMRoute 是一种混合多播协议，它首先建立核心网格，然后通过核心网格节点去建立多播树。当网络拓扑变化时，只要网格节点与树节点之间还存在链路就不需要对多播树进行更新。其缺点是由于节点的移动，可能导致路由环路和非优化的多播树。

1.2.4　服务质量保证

无线自组织网络一方面作为一种自治系统，有自身特殊的路由协议和网络管理机制；另一方面作为互联网在无线和移动范畴的扩展和延伸，它又必须能够提供到互联网的无缝的接入机制。当前互联网已经可以在一定程度上保证综合业务传输的服务质量(QoS)。近年来随着多媒体应用的普及和无线自组织网络在商业应用的进展，人们很自然地会产生在 Ad hoc 网络上传送综合业务的需求，并且希望能像固定的有线网络一样为不同业务的服务

质量提供保障。因此无线自组织网络对 QoS 保障的支持显得越来越迫切和重要。但是与固定的有线网络不同，在 Ad hoc 网络中提供 QoS 支持将面临许多不同于传统网络的新问题和挑战。

QoS 路由的目的是为应用服务寻找满足其 QoS 要求、具有足够网络资源的“端到端”传输路径。由于存在拓扑结构动态变化、链路资源时变、节点资源有限及运动形态不确定等问题。Ad hoc 网络在为路由协议提供精确的链路状态信息，为路由计算承担所需的大量资源开销、路由维护等方面较为困难。因此，QoS 路由技术一直是一个研究热点问题，也是 IETF 工作组关注的重点。与传统路由算法的分类方法相似 QoS 路由通常可分为预先式路由和按需式路由两大类。其中预先式 QoS 路由协议要求每个节点维持一至多张表以存储链路状态信息或基本路由信息，并通过广播方式进行信息更新。当需要进行 QoS 路由计算时，源节点或者根据信息表的有关内容按 QoS 参数约束直接计算出所需的可行路径，或者确定出基本路由策略，并采用部分洪泛机制发送带有 QoS 参数要求的路由请求分组，由中间节点根据其信息表的有关内容分布式完成寻路工作。相关的预先式 QoS 路由协议有 CEDAR[52]、TBP[53]、基于带宽的 QoS 路由协议[54]等。按需式 QoS 路由协议指当有分组传送需求时才进行 QoS 路由计算的寻路方式。开始传送业务分组前，源节点触发一个路由寻找进程，通常采用洪泛机制发送路由请求分组，并由中间节点根据其接口的基本信息和 QoS 参数要求分布式完成路由寻找工作。相关的按需式 QoS 路由协议有 AODV 扩展 QoS 路由协议[55]、R. Lin 的按需 QoS 路由协议[56]等。

预先式 QoS 路由协议对网络状态信息维护和更新有较高的要求，并由此造成有限网络资源的巨大浪费，而按需驱动式 QoS 路由协议，由业务需求激发往往造成实时业务分组传输的延迟和停顿。因此近年来这两类 QoS 路由协议正逐渐走向融合。将 MAC 层的多址技术和网络层路由技术相结合，通过资源预留来满足 QoS 寻路要求，正成为 QoS 路由技术研究领域的新思路。此外，多通道路由多播、路由自适应服务优先 QoS 路由、功率路由、辅助位置路由等也将是 QoS 路由技术未来的研究方向。

1.2.5 网络管理

在任何网络的建设中，控制网络、使网络具有最高效率和可靠工作，网络管理将是一个必须内容，这一过程通常包括数据收集、数据处理、数据分析和网管动作等。为了实现对网络的控制和管理，OSI 将网络管理划分为五大功能域，这就是我们通常所说的配置管理、性能管理、安全管理、计费管理和故障管理，然而由于 Ad hoc 网络的特性决定了管理上比有线网络复杂许多，因为网络拓扑的动态变化，要求网络管理也是动态自动配置。而且要考虑到移动节点本身的限制，例如能源有限、链路状态变化和有限的存储能力等，因此，要将管理协议给整个网络带来的负荷考虑在内。最后还要考虑到网络管理对不同环境的适用性等。

现阶段，就无线自组织网络的管理研究刚刚处于起步，由于网络特性排斥集中和完全分布的管理体系，因此考虑到网络信息开销和节点的移动性，一种作为最好折中的三级层次结构成为首选，最后一级由被称为代理(agent)的被管节点组成，许多互相邻近的代理集结成簇 Cluster，由第二级被称为簇首 Cluster head 的节点来管理，而簇首又由被称为网络管理者 Manager 的节点来管理[57]。

1.2.6 网络安全

与固定有线网络相比,无线自组织网络面临更多的安全威胁。在固定网络中,入侵者需搭接电缆才能偷听,需要寻找防火墙或网关的漏洞才能访问内部资源。但对于无线自组织网络,无线信道使得窃听随处可在,节点的移动性使得防火墙无法应用。无线自组织网络比固定网络更容易遭受各种安全的威胁,如窃听、伪造身份、重放、篡改报文和拒绝服务等。因此,无线自组织网络更需要安全的保障。

1.3 无线自组织网络的研究机构及研究方向

目前,无线自组织网络是网络和通信方面一个重要的研究方向,许多大学和研究机构都建立相关的研究组和实验室对无线自组织网络进行研究。

1996 年 IETF 成立了 MANET 工作组[10],目前的核心任务是设计无线自组织网络的路由协议和进行协议标准化工作。MANET 工作组已经提出了许多协议草案,如:DSR[30]、AODV[31]、OLSR[28]、TBRPF[29]等。

Columbia 大学的 INSIGNIA 项目[58],对无线自组织网络环境下支持 QoS 路由的框架进行了设计、实现和评估。主要研究方向为无线自组织网络路由的 QoS 保障。

Rice 大学的 Monarch 项目组[59]对移动网络通信架构进行研究,其目标是使移动节点在任何时间、任何地点与移动节点、固定节点进行无缝的通信。研究范围包括:无线自组织网络路由协议的设计与实现、性能评估、服务质量保证、安全协议、多播等。

Cornell 大学的无线网络实验室(Wireless Network Lab)[60]致力于无线通信协议、算法的研究。在无线自组织网络方向,其研究范围包括:路由协议、资源管理、MAC 层协议、QoS、安全算法等。

Maryland 大学的移动计算与多媒体实验室(The Mobile Computing and Multimedia Laboratory)[61],研究移动计算和多媒体系统。其研究范围包括无线自组织网络的路由协议、QoS 等。

Stanford 大学的移动计算研究组(The Mobile Computing Group)[62],其研究方向为无线自组织网络节点之间增强合作的算法。

加州大学洛杉矶分校(UCLA)的无线自适应移动实验室(Wireless Adaptive Mobility Laboratory)[63],其研究重点为:无线自组织网络环境下协议的设计、实现和性能测试。网络模拟实验平台的研究与设计。

Illinois 大学的移动环境实验室(The Illinois Mobile Environments Laboratory)[64],致力于无线网络,特别是无线自组织网络 QoS 的系统架构、协议和中间件的研究。

Rutgers 大学的无线信息网络实验室(Wireless Information Network Laboratory)[65],研究方向为无线自组织网络路由安全协议。

瑞士联邦支持的 Terminode 项目[66],旨在研究和模拟大规模、自组织的无线自组织网络,其研究领域包括从物理导到应用层的所有层面,以层间的相互作用,如路由算法、移动性管理、节点相互协作的激励机制和安全策略等。

加州大学圣芭芭拉分校(UCSB)的移动管理与网络实验室(Mobility Management and

Networking Laboratory)[67],研究方向为路由协议、多播、QoS 协议等。

参考文献

[1] William Stallings. Wireless Communications and Networks. Prentice Hall,2002.

[2] Johnson D B,Maltz D A. Protocols for Adaptive Wireless and Mobile Networking. IEEE Personal Communications Magazine,February 1996,34-41.

[3] Ayanoglu E,Eng K Y,Karol M J. Wireless ATM: Limits,Challenges,and Proposals,IEEE Personal Communications Magazine,August 1996,18-34.

[4] Jubin J,JD Tornow. The DARPA Packet Radio Network Protocols,in Proceedings of the IEEE,Special Issue on Packet Radio Networks,75(1),January 1987,21-32.

[5] Corson S,Macker J,Mobile Ad hoc Networking(MANET): Routing Protocol Performance Issues and Evaluation Considerations,RFC 2501,January 1999.

[6] Beyer D A. Accomplishments of the DARPA Survivable Adaptive Networks SURAN Program. In Proceedings of the IEEE MILCOM Conference,1990.

[7] Leiner B M,Robert Ruth,Sastry A R. Goals and Challenges of the DARPA GloMo Program. IEEE Personal Communications,1996,3(6): 34-43.

[8] The Joint Tactical Radio System,http://jtrs. army. mil.

[9] http://www. ieee802. org/11.

[10] MANET working group,http://www. ietf. org/html. charters/manet-charter. html.

[11] Finke C R. TPRS Quarterly Report. Texas Packet Radio Society,Feb. 1992.

[12] Chen Z D,Kung H T,Vlah D. Ad hoc Relay Wireless Networks over Moving Vehicles on Highways. ACM Symposium on Mobihoc,California,Oct 2001.

[13] George N,Rahim T. On the Relaying capability of Next-Generation GSM Cellular Networks. IEEE Personal Commun Mag,Feb 2001: 40-47.

[14] Ljubica B,Levente B,Srdjan C,et al. Self-organization in mobile Ad hoc network: the approach of terminodes. IEEE Communication Magazine,2001,39(6): 166-174.

[15] Fullmer C L,Garcia-Luna-Aceves J J. Solutions to hidden terminal problems in wireless networks,Proceeding ACM SIGCOM'97,Cannes,France,1997: 39-49.

[16] Lichun B,Garcia-Luna-Aceves J J. Collision-free topology-dependent channel access scheduling,MILCOM2000,Los Angeles,Califomia,2000: 507-511.

[17] Royer E M,Lee S J,Perkins C E. The Effects of MAC Protocols on Ad hoc Communication. IEEE Wireless Communication and Networking Conference (WCNC 2000),Chicago,Sept. 2000.

[18] Karn P,MACA—A new channel access method for packet radio,ARRL/CRRL Amateur 9th computer networking conference,April 1990: 134-140.

[19] Bharghavan V,Demers A,Shenker S,et al. MACAW: A Media Access Protocol for Wireless LAN,in proceeding of ACM SIGCOMM'94,1994: 212-225.

[20] Fullmer C L,Garcia-Luna-Aceves J J. Floor Acquisition Multiple Access (FAMA) For Packet-Radio Networks,In Proceedings of the ACM SIGCOMM'95,Sep 1995: 262-273.

[21] Haas Z J,Deng J. Dual Busy Tone Multiple Access(DBTMA)-Performance Evaluation,IEEE Semiannual Vehicular Technology Conference (VTC'99),Houston,TX,May 1999.

[22] Hong X,Xu K,Gerla M. Scalable routing protocols for mobile ad hoc networks,IEEE Network,2002,16 (4): 11-21.

[23] Mauve M,Widmer J,Hartenstein H. A Survey on Position-Based Routing in Mobile Ad hoc Networks. IEEE Network,2001,1(6): 30-39.

[24] Josh Broch, Maltz D A, Johnson D B, et al. A Performance Comparison of Multi-Hop Wireless Ad hoc Network Routing Protocols. Proceedings of the Fourth Annual ACM/IEEE International conference on Mobile computing and networking (MobiCom'98), Dallas, 1998, October 25-30.

[25] Royer E M, Toh C K. A Review of Current Routing Protocols for Ad-hoc Mobile Networks, IEEE Personal Communications, 1999, 6(2): 46-55.

[26] Pei G, Gerla M, Chen T W. Fisheye State Routing: A Routing Scheme for Ad hoc Wireless Networks, Proceeding in IEEE International Conference on Communications (ICC2000), New Orleans, June 2000.

[27] Perkins C E, Bhagwat P. Highly Dynamic Destination Sequenced Distance-vector Routing (DSDV) for mobile computers, The ACM SIGCOMM Conference on Communications Architectures, London, 1994.

[28] Clausen T, Jacquet P. Optimized Link State Routing Protocol (OLSR), RFC3626, October, 2003.

[29] Ogier R, Templin F, Lewis M. Topology Dissemination Based on Reverse-Path Forwarding (TBRPF), RFC3648, February, 2004.

[30] Johnson D B, Maltz D A, Hu Y C. The Dynamic Source Routing Protocol for Mobile Ad hoc Networks (DSR), INTERNET-DRAFT, draft-ietf-manet-dsr-10. txt, 19 July 2004.

[31] Perkins C, Belding-Royer E, Das S. Ad hoc On-Demand Distance Vector (AODV) Routing, RFC3561, July 2003.

[32] Pei G et al. A Wireless Hierarchical Routing Protocol with Group Mobility, IEEE Wireless Communications and Networking Conference (WCNC'99), New Orleans, Sept. 1999.

[33] Chiang C C, Gerla M. Routing and Multicast in Multihop, Mobile Wireless Networks, Proceedings of IEEE International Conference on Universal Personal Communications (ICUPC'97), San Diego, USA, Oct. 1997.

[34] Gerla M, Hong X, Pei G. Landmark Routing for Large Ad hoc Wireless Networks, Proceedings of IEEE Global Communications Conference (GLOBECOM 2000), San Francisco, USA, Nov. 2000.

[35] Haas Z J, Pearlman M R. The Performance of Query Control Schemes for the Zone Routing Protocol, The IEEE/ACM Transactions on Networking, 2001, 9(4): 427-438.

[36] Navas J C, Imielinski T. Geographic Addressing and Routing, In Proceeding 3th ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'97), Budapest, Hungary, Sept. 1997: 26-30.

[37] Basagni S et al. A Distance Routing Effect Algorithm for Mobility (DREAM), In Proceeding 4th ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'98), Dallas, Oct. 1998: 76-84.

[38] Ko Y B, Vaidya N H. Location-aided Routing(LAR) in Mobile Ad hoc Networks, In Proceeding 4th ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom'98), Dallas, Oct. 1998: 66-75.

[39] Karp B, Kung H T. GPSR: Greedy Perimeter Stateless Routing for Wireless Networks, In Prodeeding 6th ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom2000), Boston, 2000: 243-254.

[40] Gomez-Castellanos J, Campbell A T, Naghshineh M, C. Bisdikian PARO: Supporting Transmission Power Controlled Routing in Wireless Ad hoc Networks, ACM/Kluwer Journal on Wireless Networks (WINET), 2003, 9(5): 443-460.

[41] Singh S, Raghavendra C S. PAMAS-Power Aware Multi-Access protocol with Signalling for Ad hoc Networks. ACM Computer Communication Review, July 1998, 5-26.

[42] Cordeiro C, Gossain H, Agrawal D. Multicast over Wireless Mobile Ad hoc Networks: Present and

Future Directions. IEEE Network Magazine,January/February 2003,17(1): 52-59.
[43] Wu C W,Tay Y C,Toh C K. Ad hoc Multicast Routing Protocol Utilizing Increasing id-numberS (AMRIS) Functional Specification. Internet draft, Nov. 1998. http://www. ietf. org/proceedings/99nov/I-D/draft-ietf-manet-amris-spec-00. txt.
[44] Royer E M, Perkins C E. Multicast Operation of the Ad hoc On-Demand Distance Vector Routing Protocol, in Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MOBICOM'99),Aug. 1999: 207-218.
[45] Jiand L,Scott Corson M. A Lightweight Adaptive Multicast Algorithm,Proceedings of IEEE Global Communications Conference(GLOBECOM 1998),1036-1042.
[46] Chen K, Nahrstedt K. Effective Location-Guided Tree Construction Algorithms for Small Group Multicast in MANET,The 21st Annual Joint Conference of the IEEE Computer and Communications (INFOCOMM2002),New York,June,2002: 1180-1189.
[47] Gerla M,Lee S J,Su W. On-Demand Multicast Routing Protocol (ODMRP) for Ad hoc Networks, Internet draft,draft-ietf-manet-odmrp-02. txt,2000.
[48] Garcia-Luna-Aceves J J,Madruga E L. The Core-Assisted Mesh Protocol,IEEE Journal on Selected Areas in Communications,Special Issue on Ad-hoc Networks,1999,17(8): 1380-1394.
[49] Chiang C C,Gerla M,Zhang L. Forwarding Group Multicast Protocol (FGMP) for Multihop,Mobile Wireless Networks,ACM-Baltzer Journal of Cluster Computing: Special Issue on Mobile Computing, 1998,1(2): 187-196.
[50] Bommaiah E et al. AMRoute: Adhoc Multicast Routing Protocol,Internet draft,Aug. 1998.
[51] Sinha P, Sivakumar R, Bharghavan V. MCEDAR: Multicast Core-Extraction Distributed Ad hoc Routing,IEEE Wireless Communications and Networking Conference(WCNC1999), Sept. 1999: 1313-1317.
[52] Sivakumar R, Sinha P, Bharghavan V. CEDAR: A Core-extraction Distributed Ad hoc Routing Algorithm,IEEE Journal on Selected Areas in Communications,1999,17(8): 1454-1465.
[53] Chen S,Nahrstedt K. Distributed Quality-of-Service Routing in Ad hoc Networks,IEEE Journal on Selected Areas in Communications,1999,17(8): 1488-1505.
[54] Lin R, Liu J S. QoS routing in ad hoc wireless networks, IEEE Journal on Selected Areas in Communications,1999,17(8): 1426-1438.
[55] Perkins C E, Royer E M, Das S R. Quality of Service for Ad hoc On-Demand Distance Vector Routing,Internet draft,14 October 2003,http://people. nokia. net/charliep/txt/aodvid/qos. txt.
[56] Lin R. On-demand QoS Routing in Multihop Mobile Networks, Proc. INFOCOM' 2001, 3: 1735-1744.
[57] Chen W,Jain N,Singh S. ANMP Ad hoc Network Management Protocol,IEEE Journal On Selected Areas In Communications,1999,17(8).
[58] The INSIGNIA QoS Framework,http://comet. columbia. edu/insignia/overview. html.
[59] The Rice University Monarch Project,http://www. monarch. cs. cmu. edu.
[60] The Wireless Networks Laboratory at Cornell University, http://people. ece. cornell. edu/~haas/wnl.
[61] The Mobile Computing and Multimedia Laboratory University of Maryland,College Park,http://www. cs. umd. edu/projects/mcml/index. html.
[62] The Mobile Computing Group at Stanford University, http://mosquitonet. stanford. edu/index. html.
[63] The UCLA Wireless Adaptive Mobility Laboratory,http://www. cs. ucla. edu/NRL/wireless.
[64] The Illinois Mobile Environments Laboratory,http://timely. crhc. uiuc. edu/index. html.

[65] Wireless Information Network Laboratory, http://www.winlab.rutgers.edu/pub/Index.html.

[66] Hubaux J P, Gross T, Le Boudec J Y, et al. Towards self-organized mobile Ad hoc Networks: The Terminodes Project, IEEE Communications Magazine, January 2001.

[67] Mobility Management and Networking (MOMENT) Laboratory, http://moment.cs.ucsb.edu.

[68] The 23rd Conference of the IEEE Communications Society, http://www.ieee-infocom.org/2004/index.html.

[69] The fifth ACM International Symposium on Mobile Ad hoc Networking and Computing, http://www.sigmobile.org/mobihoc/2004/cfp.html.

[70] 2nd Annual IEEE Communications Society Conference on Sensor and Ad hoc Communications and Networks, http://www.ieee-secon.org/2005.

CHAPTER 2

第2章　无线自组织网络安全的研究进展

摘要：无线自组织网络是一种完全由移动主机构成的网络，其主要特点为网络拓扑易变，带宽、能源有限。这些特点对于无线自组织网络安全的设计与实现提出了巨大的挑战。本章介绍了无线自组织网络安全研究的最新研究进展，分为密钥管理、路由安全、入侵检测、增强合作几个方面，对一些典型方案进行了说明，分析了各种方案的优点和缺点，并进行了综合比较。文中分析了目前协议存在的一些问题并提出了相应的改进方法，最后指出了下一步研究方向。

关键字：计算机网络、信息安全、综述、无线自组织网络、密钥管理、路由安全、入侵检测、增强合作。

2.1　引言

无线自组织网络作为一种新型的移动多跳无线网络，与传统的无线网络有很大不同，它不依赖于任何固定的基础设施和管理中心，而是通过传输范围有限的移动节点间的相互协作和自我组织来保持网络连接和实现数据的传递。无线自组织网络的独特的结构，从而产生下列一些突出的特点[1]。

(1) 动态的拓扑结构：节点可在网络中任意移动，随时加入和退出网络。

(2) 有限的资源：无线通信带宽有限，移动节点的能源也有限。

(3) 多跳的通信：无线节点发射功率有限，发送报文到接收区域外的节点时，需要其他节点来中转信息。因此，任意一个节点既是主机又是路由器。

(4) 脆弱的网络安全：由网络的自组织性、节点的移动性和无线通信信道，使得无线自组织网络更容易遭受各种攻击，其安全问题更加严峻。

无线自组织网络最初用于军事领域，如战场上坦克之间和海面上舰艇之间的组网，但是由于其建网方式灵活、配置快捷方便，构造成本较低等优势，它逐渐运用于商业和民用环境之中，如会议数据交换、紧急援救、偏远地区等一些需要临时组网的应用中。

与固定有线网络相比，无线自组织网络面临更多的安全威胁。在固定网络中，窃听者需搭接电缆才能偷听，需要寻找防火墙或网关的漏洞才能访问内部资源。但对于无线自组织网络，无线信道使得窃听随处可在，节点的移动性使得敌我双方无边界，防火墙无法应用。因此，无线自组织网络比固定

网络更容易遭受各种安全的威胁，如窃听、伪造身份、重放、篡改报文和拒绝服务等。

本章首先分析了无线自组织网络的安全弱点和安全目标，然后综述了现行的各种解决方法，并指出了各种方案的优点和缺点。本章安排如下：2.2 节介绍安全弱点，主要分析由无线自组织网络独特的结构所造成在安全方面的弱点及安全所要求达到的目标。2.3 节密钥管理，论述了在无线自组织网络环境下如何实现密钥的分配与管理。2.4 节路由的安全，首先分析路由协议的威胁，其次介绍几种典型的路由安全方案。2.5 节讨论入侵检测，介绍了入侵检测的系统结构和检测模式。2.6 节增强节点合作，介绍了两种促使节点参与网络交换的机制。2.7 节总结和展望，对全文进行了总结，提出了下一个研究方向。

2.2　无线自组织网络的安全弱点和安全目标

2.2.1　安全弱点

传统网络中，主机之间的连接是固定的，网络采用层次化的体系结构，并具有稳定的拓扑。传统网络提供了多种服务以充分利用网络的现有资源，包括路由器服务、命名服务、目录服务等，并且在此基础上实现了相关的安全策略，如加密、认证、访问控制和权限管理、防火墙等。而在无线自组织网络中没有基站或中心节点，所有节点都是移动的，网络的拓扑结构动态变化。并且节点间通过无线信道相连，没有专门的路由器，节点自身同时需要充当路由器，也没有命名服务、目录服务等网络功能。两者的区别导致了在传统网络中能够较好工作的安全机制不再适用于无线自组织网络，主要表现在以下几个方面。

1. 传输信道方面

无线自组织网络采用无线信号作为传输媒介，其信息在空中传输，无需像有线网络一样，要切割通信电缆并搭接才能偷听，任何人都可接收，所以容易被敌方窃听。无线信道又容易遭受敌方的干扰与注入假报文。

2. 移动节点方面

因为节点是自主移动的，不像固定网络节点可以放在安全的房间内，特别是当无线自组织网络布置于战场时，其节点本身的安全性是十分脆弱的。节点移动时可能落入敌手而投降，节点内的密钥、报文等信息都会被破获，投降后的节点又可能以正常的面目重新加入网络，用来获取秘密和破坏网络的正常功能。因此，无线自组织网络不仅要防范外部的入侵，而且要对付内部投降节点的攻击。

3. 动态的拓扑

无线自组织网络中节点的位置是不固定的，可随时移动，造成网络的拓扑不断变化。一条正确的路由可能由于目的节点移动到通信范围之外而不可达，也可能由于路由途经的中间节点移走而中断。因此，难于区别一条错误的路由是因为节点是移动造成的还是虚假路由信息形成的。由于节点的移动性，在某处被识别的恶意节点移动到新的地点，改变标识后，它可重新加入网络。另外由于动态的拓扑，网络没有边界，防火墙也无法应用。

4. 安全机制方面

在传统的公钥密码体制中,用户采用加密、数字签名、报文鉴别码等技术来实现信息的机密性、完整性、不可抵赖性等安全服务。然而它需要一个信任的认证中心来提供密钥管理服务。但在无线自组织网络中不允许存在单一的认证中心,否则不仅单个认证中心的崩溃将会造成整个网络无法获得认证,而且更为严重的是,被攻破认证中心的私钥可能会泄露给攻击者,攻击者可以使用其私钥来签发错误的证书,假冒网络中任意一个移动节点,或废除所有合法的证书,致使网络完全失去了安全性。若通过备份认证中心的方法虽然提高了抗毁性,但也增加了被攻击的目标,任意一个认证中心被攻破,则整个网络就失去了安全性。

5. 路由协议方面

路由协议的实现也是一个安全的弱点,路由算法都假定网络中所有节点是相互合作的,共同去完成网络信息的传递。如果某些节点为节省本身的资源而停止转发数据,这就会影响整个网络性能。更可怕的是投降节点和参与到网络中的恶意节点专门广播假的路由信息,或故意散布大量的无用数据包,从而导致整个网络的崩溃。

2.2.2 安全目标

1. 机密性

机密性是保证相关的信息不会泄露给未授权的用户和实体,通常采用加密的方法来实现。由于无线自组织网络采用的是无线信道,更容易受到窃听攻击,所以在网络上传输的敏感信息,必须机密可靠,否则这些信息被敌方破获,后果将不堪设想。

2. 完整性

完整性用于保证信息在传输过程中不被篡改,确保收到的消息与发送的消息一样,没有冗余、插入、修改。通常采用报文摘要或散列算法来实现。对路由信息的恶意篡改会造成回路、网络分裂,所以也需进行完整性保护。

3. 安全认证

使每个节点能够确认与其通信的节点的身份。如果没有认证,攻击者就可以假冒网络中的节点来与其他节点进行通信,并可以获得那些未被授权的资源和敏感信息,并以此威胁整个网络的安全。

4. 不可抵赖

不可抵赖是防止发送方抵赖所传输的消息。因此,当发送一个报文时,接收方能够证实该消息确实是由所宣称的发送方发送的。这可用来发现和孤立恶意节点。

5. 可用性

可用性确保网络节点在受到各种网络攻击时仍然能够提供相应的服务。在 Ad hoc 网络中拒绝服务攻击可以在各个协议层进行,在物理层和链路层,攻击者可以通过无线干扰来扰乱通信信道,在网络层,攻击者可以破坏路由信息,使网络无法互联,在高层,攻击者可以通过伪造各种应用使高层服务紊乱。

2.3　密钥管理

由于无线自组织网络具有自组织和动态拓扑的特性，使得在固定网络中常用的密钥管理手段无法在 Ad hoc 网络中应用，例如：Certification Authority(CA)或 Key Distribution Center(KDC)就无法在无线自组织网络应用，使用这些设施其一容易导致单点失败和拒绝服务，即该设施由于敌方攻击而失灵了，整个网络就不能正常运转了；其二由于无线多跳通信误码率高和网络拓扑动态变化，会大大降低服务的成功率，延长服务时间；其三，容易导致网络拥塞，本来就不充足的传输带宽，网络中各节点还都要到该节点去认证。文献[2]模拟试验了集中认证、分布认证、本地认证三种方法的可扩展性、健壮性和有效性，集中认证性能最差，特别当网络节点数量增加、网络负载上升时，集中认证的性能急剧下降。当网络节点数量由 40 增加到 100 时，集中认证成功率由 92%降到 22%，本地认证保持在 96%，当网络负载由 0 增至 100 包/秒(每包 512 字节)时，集中认证成功率由 80%降到 45%，本地认证保持在 95%，当信道误码率从 0 增至 10%时，集中认证成功率由 80%降到 50%，本地认证保持在 93%。通过实验证实了 CA 的方法在无线自组织网络中无法应用，但是近来提出的许多 Ad hoc 路由安全协议，如 Ariadne[3]、SRP[4]等，都要求事先存在或预先分配共享密钥或公开密钥，这就要求提供适应于无线自组织网络的密钥管理手段。下面首先介绍两种具有代表性的密钥管理方案，其次介绍其他几种解决的方案。

2.3.1　自组织的密钥管理

该算法在文献[5]中首先提出并进行了概要介绍，在文献[6]中进行了详细论述。该算法不需要公认的 CA 来发布证书，节点自己发布并维护证书，类似于 PGP 算法[7]，用户根据需求拥有对方的证书来进行认证。与 PGP 算法不同的是，用户证书是分布存储于每个用户自身而不是存储于认证服务器之中。在该算法中，每个用户在本地维护一个证书数据库，当两个用户需要相互认证时，他们合并他们各自拥有的证书数据库形成一张认证路径图，并努力从该图中发现一条认证链路。如果发现一条认证路径，则认证成功，否则认证失败。图 2-1 显示合并节点 u 和节点 v 的证书数据库形成的认证图，并找到一条从节点 u 到节点 v 的认证路径。

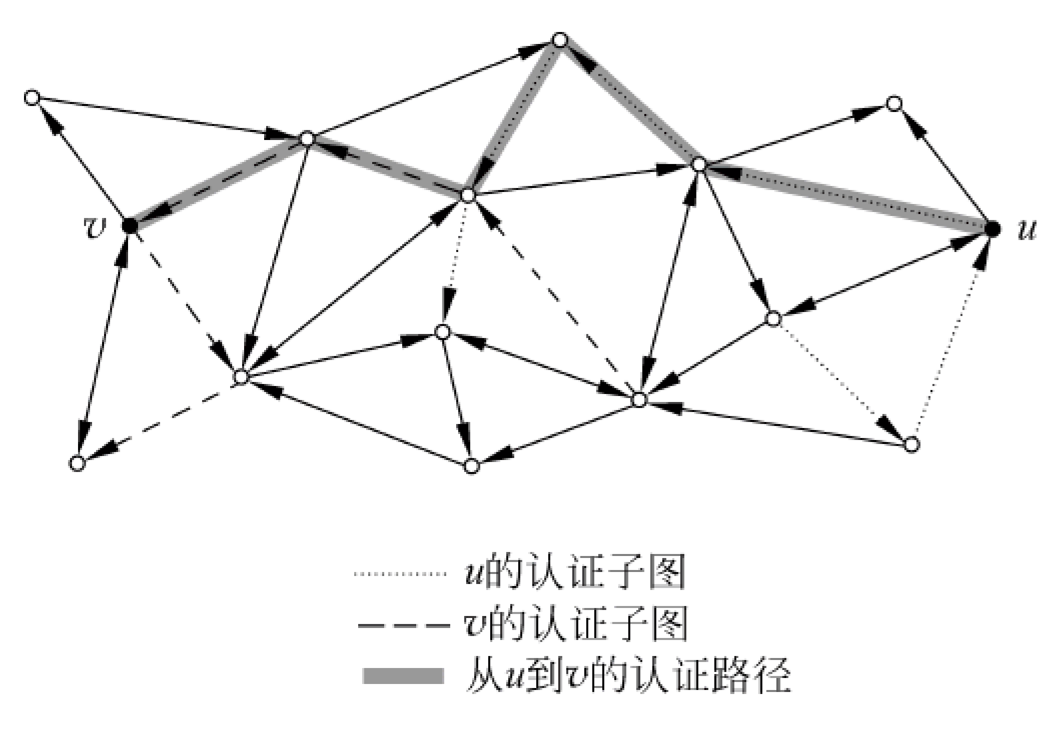


图 2-1　从节点 u 到节点 v 的认证图

认证的成功取决于用户本地证书数据库的构成方式和认证路径图的特性。所谓认证路径图就是一张节点代表用户的公钥，边代表用户颁发的公钥证书的用于节点认证的有向图。在文章中作者提出了几种用户证书数据库的构成算法并分析了它们各自在认证中表现出的性能。一种是最多认证算法，每个用户从整个认证图中选择几条路径构成自己的本地证书数据库，每条路径都以用户为起始点，并且这几条路径上的节点最多，即能够认证的节点最多。另一种是最短认证算法，构成方法是选择到某节点最短的认证路径。对两种算法的分析表明和模拟试验表明，即使本地证书数据库小到只有$\sqrt{n}$（n 为全部用户数目）时，任何两个用户仍有 90%的概率找到一条相互认证的路径。

本算法的优点在于完全不需要 CA 来发布和维护证书，防止了单点失败。缺点有三点，其一，由于没有 CA 来验证身份，任何能发布证书的节点均能加入网络，恶意节点可顶替尚未进入网络的节点或编造节点标识发布证书加入网络。其二因为节点存储证书信息不完全，不能保证 100%认证，其认证成功率与证书数据库形成密切相关。其三算法的扩展性不好，当网络扩大时，证书数据库的形成、维护和认证的花费会急骤增加。

2.3.2 分布式的密钥管理

Lidong Zhou 和 Zygmunt J. Haas 提出一种基于门限密码理论[8]，实现分布式的 CA 来进行密钥管理的算法[9]。所谓 nt 门限密码，即 n 个具有密码学操作(如数字签名)能力的实体，其中任意 t 个实体联合起来即可完成该项操作，而任意小于 t 个的实体不能执行该操作。CA 的系统公钥为每个节点所知，用来校验证书，系统私钥用来签发证书。该算法利用门限密码的特点，网络初始化时，由集中的 CA 将网络系统私钥分为 n 份，指定 n 个节点拥有，这 n 个节点就充当了分布式的 CA。当需要 CA 来发布证书时，这 n 个节点中的任意 t 个合作生成一份有效的证书。当新节点加入网络时，就可向这 n 个节点中的任意 t 个节点提出证书申请，每个节点返回部分签名证书，合起来就形成了一份完整的证书。为了防止移动攻击者(mobile adversaries)[10]攻击，即敌方攻破一个节点后转向下个节点，这样经过一段时间后会攻破很多节点，甚至达到 t 节点，采取共享更新算法，即从老的 n 份私钥中生成新的 n 份私钥。因为新的私钥独立于老私钥，所以只要更新周期合适，就能对抗 mobile adversaries 攻击。该算法的优点是将单一的 CA 服务分散到 n 个节点中去，防止了单点失败，提高了网络的健壮性，只要被攻陷节点少于 t 个，整个网络仍然是安全的。缺点是增加了网络传输负载和延长了服务时间，因为在集中式的 CA 中，节点只需与一个 CA 联系返回一份证书，而分布式 CA，节点需要到 t 个持有系统私钥的节点去申请，返回 t 份证书，而这些节点可能遍布于网络各处，需要多跳通信才能达到，只有与这 t 个节点都成功通信时，才能完成。

Jiejun Kong 等人提出了一个类似的方案[11]，改进了上述方案的缺点。网络系统初始化时，由集中的 CA 将系统私钥分为 t 份授权给 t 个节点，然后由这 t 个节点联合起来继续将私钥授予网络中其余各节点。这样，系统私钥不只分为 n 份由 n 个节点持有，而是网络中每个节点都持有一份系统私钥。节点加入网络时，只要向周围 t 个邻居节点提出申请，即可获得证书。该方案提高可用性，降低了网络负载。该方案还具有扩展性，无论网络扩大或缩小都可适用。文献[12]提出的方案与文献[11]基本相同，只是增加了邻居监视功能，每个节

点的周围的邻居不仅联合颁发证书，而且负责监视其的行为。如果该节点有恶意行为，证书到期后将不能申请到证书。

Aram Khalili 等人提出了一种基于用户标识的门限密码管理算法[13]，它将用户标识作为节点的公钥，节点的私钥由 t 个持有系统私钥的节点联合签发。它的优点在于节点不再需要分发公钥，进一步降低了网络负载和节点计算量。上述方案有个共同的弱点，无法对抗 Sybil 攻击[14]，即攻击者利用尽可能多的节点标识来获得私钥共享份数，最终可以重建系统私钥。

2.3.3　两种密钥管理方案的比较和分析

分布式和自组织的密钥管理是较为典型的两种密钥管理方案，表 2-1 对这两种方案进行了比较。上述各种密钥管理方案都考虑到无线自组织网络自组织无中心的特点，设计上采取各种方法代替集中的 CA，如采用分布式的 CA、自组织的方法，但同时也导致了一些缺陷。

表 2-1　分布式密钥管理与自组织密钥管理的比较

算法分类	分布式的密钥管理	自组织的密钥管理
理论基础	门限密码	PGP 技术
前提条件	信任实体生成系统公钥和私钥，将公钥发送给所有节点，将私钥分为 n 份授予 n 个节点持有	各节点通过证书交换，形成本地证书数据库
证书管理	n 个持有系统私钥的节点中 k 个联合颁发并管理	由节点自己生成并管理
认证方式	通过系统公钥校验节点的证书	通过证书链
优势	防止了单点失败	防止了单点失败
劣势	增加了计算负载和网络流量	攻击者可假冒合法节点发布证书加入网络、算法的扩展性不好

密钥管理方案存在问题及改进方法：

(1) 在分布式的密钥管理中，将集中的 CA 分配到 n 个节点中去，由 n 个节点中的 k 个节点合作签发证书。节点必须与 k 个节点都建立通信并成功申请到 k 份证书才能合成一份有效的证书。这 k 个节点可能分布在网络各处，需要多跳通信才能到达。如果与 k 个节点中任意一个节点通信失败或返回证书有误，则无法合成证书，整个申请失败，必须再找 k 个节点重新开始申请。从上述过程可以看出，分布式的 CA 虽然防止了单点失败，但也增加了网络负载，延长了服务时间，降低了申请证书的成功率。解决该问题的方法有两种，一种是减少节点需要申请证书的数目 k。极端情况下 $k=1$，任何一个拥有系统私钥的节点均可签证书。这样就减少了需要建立通信的节点数目。该方法虽然降低了网络负载，但也降低了系统的安全性，攻击者需要攻破的节点的数目也减少了。另一种是增加拥有系统私钥的节点数目 n。极端情况下 n 为网络中的节点数，即网络中所有节点都持有一份系统私钥，这种情况下节点只要向周围 k 个邻居节点申请证书即可。由于通信局限于本地，网络负载和服务时间会大大降低，但随着拥有系统私钥的节点数目的增加，对其管理和维护的费用也在增加，如系统私钥更新的费用就会明显增加。权衡利弊，最好采用自适应的方式来决定 n 和 k 的数目。当网络处于相对安全的环境下运行时，则可降低 k 的数目，反之则提高 k 的数目。

当网络节点数目多时,N 值可取大些,反之则可取小些。

(2) 在分布式的密钥管理方案中,还存在个共同的弱点,无法对抗 Sybil 攻击[14],即攻击者利用尽可能多的节点标识来获得私钥共享份数,最终可以重建系统私钥。对抗 Sybil 攻击可综合采用以下方法。首先绑定节点硬件地址与标识,使攻击者不能随意改变其标识。其次,及时更新系统私钥,使攻击者不能取得 k 份系统私钥。最后采用入侵检测的方法,监视节点行为,发现不良行为,及时予以告警。

(3) 在自组织的密钥管理方案中,每个节点负责颁发和维护自己的证书,由于没有 CA 来验证身份,任何能发布证书的节点均能加入网络,攻击者可假冒或编造节点标识发布证书加入网络。解决该问题可采用,首先通过检测证书的一致性来发现假冒行为,如同一份证书代表了两个节点,然后通过邻居监视来确定假冒者,从而将其排除出网络。

2.3.4 其他一些密钥管理方案

Frank Stajano 和 Ross Anderson 提出复活鸭子的安全模式[15],鸭子破壳而出之后,它会把它见到的第一个移动物体作为它的母亲。与此类似,节点初始化时,它将第一个发给它密钥的节点作为它的拥有者,它只接受拥有者的控制。这种控制一直保持到节点死亡,节点重新复活后可产生新的拥有者。这样就形成了一种树状的密钥分发与管理模式。这种方案适用于大型层次型的网络,例如,战场指挥网络。它还适用于低价的嵌入式设备,如传感器网络(sensor networks)[16]。它的优点在于简单,不需复杂的计算。缺点是缺乏灵活性,如果一个节点失灵了,它所带的所有子节点和孙节点都将无法进行安全通信。

N. Asokan 和 Philip Ginzboorg 提出一种基于口令的密钥管理方案[17]。该方案针对一群带着笔记本电脑的人在一间会议室里开会,在没有任何安全架构的情况下,建立各移动电脑之间的安全信息交换。它的基本思想是从一个弱的口令字,通过多方 Diffie-Hellman 密钥交换[18],最终生成用于信息安全交换的密钥。该方案也不需要 CA 或 KDC,但只适用于小范围,扩展性不好。

Zheng Yan 提出基于外部 CA 的密钥管理[19]。外部 CA 可设立在卫星或飞机上,采用广播加密技术来发送信息。网络中节点嵌入专用硬件来实现密钥的存储、加密和解密操作。该方案对硬件要求过高,适用面不广。

Srdjan Capkun 提出通过节点的移动来建立交换密钥的方案[20]。它认为节点的移动性能够帮助网络安全的实现。因为无线自组织网络中节点在频繁移动中,所以任意两节点有机会相互见面,当它们接近到一定程度时,通过安全旁路(如红外信道)相互交换密钥,以此种方式实现密钥的建立和维护。该方案也不需要 CA,但受到节点移动模式、分布范围等因素的影响。

文献[21~23]提出组密钥管理方案,它们共同的特点是将网络分为多个组,组头负责组内各节点的密钥管理。该方案在一定程度上防止了单点失败,但由于节点的移动性,组头和组员管理十分复杂。

上述所有方案都考虑到无线自组织网络自组织无中心的特点,设计上采取各种方法取消或代替集中的 CA,如采用分布式的 CA 或自组织的方法,但同时也导致了一些缺陷,如:在自组织的密钥管理中,网络无法排除能够编造节点标识并发布证书的恶意节点,在分布式

的 CA 中，节点需要向多个节点去申请证书，增加网络负载，延长了服务时间。总之，各个方案都有各自的优势和劣势，适用于不同的场合，如：少数人开会可采用基于口令的密钥分配，战场指挥网络可采用复活鸭子或组密钥管理，在单个管理域内可采用分布式的 CA，在多个管理域内可采用自组织的密钥管理。

2.4　路由安全

路由协议是无线自组织网络中的一个重要的部分，因为它直接决定了网络功能的实现和效率。由于无线自组织网络与固定网络具有不同的特点，如节点移动、多变拓扑，使得常规的路由协议不适用于无线自组织网络。近年来提出了许多适用于无线自组织网络的路由协议，如 DSR[24]、AODV[25]、DSDV[26] 等，这些路由协议在设计时充分考虑了 Ad hoc 的特点，却没有考虑到安全方面的因素。这使得上述路由协议在安全方面存在重大隐患。下面首先分析路由协议的威胁，其次介绍几种典型的路由安全方案。

2.4.1　路由安全的威胁

1. 篡改

路由协议假定网络中节点都是相互合作的，转发报文的节点不会修改与其无关的路由信息，所以不检查路由信息的完整性。这使攻击者能够十分容易更改路由信息中任何字段，例如：AODV 路由中的序号和跳数，DSR 路由包中的路由节点序列等，从而产生错误的路由，如重定向、回路等，导致整个网络性能下降。攻击者能够篡改路由报文的根本原因在于节点无法对路由报文进行完整性检测。

2. 冒充

因为路由协议并不认证报文的地址，所以攻击者可以声称为某个节点加入网络，甚至能够屏蔽某个合法节点，替他接收报文。其根本原因在于节点不能鉴别报文的来源。

3. 伪造

攻击者可以伪造并广播假的路由信息。例如：广播某条存在的路由已中断，或编造一条并不存在路由。它可造成回路、分割网络、孤立节点等。其原因在于无法验证报文的内容。

4. 泄露拓扑结构

在路由查询和发送报文中都包含有明确的路由信息，如 DSR 报文头部就含有从源节点到目的节点的路由。攻击者能够通过偷听这些报文分析出节点相邻情况、所处位置等拓扑信息，可进一步通过流量分析，得出节点在网络中的功能和角色。借助这些信息，攻击者可准确地进攻网络控制节点或军事网络中的指挥员。

5. 不合作

无线自组织网络中节点既是终端用户又作为一个网络交换的路由器，它不仅收发自己的报文，还要转发其他节点的报文。一个源节点到目的节点平均有五跳的网络，转发所占的能量可达整个能量消耗的 80%[27]。因此，在多个管理域的网络中，某些自私的节点为了节省自己的资源，不参加路由查询、不转发报文，这会严重影响网络的性能。当有 10%～40%

节点出现自私行为时，网络吞吐量将下降16%～32%[28]。

6. 几种路由攻击

资源消耗攻击：攻击者发送大量无用数据报文，消耗网络和节点资源，如带宽、内存、CPU、电池等。

Wormhole攻击[29]：两个串通的攻击者，采用专用通路直接相连，越过正常的拓扑结构，直接转发路由查询报文，造成错误的路由拓扑信息。图2-2为Wormhole攻击示意图，从S节点到D节点的正常路由应该为$S—A—B—C—D$，但攻击者M_1和M_2通过ABC建立虚拟专用通道用来转发路由查询报文，这样形成了$S—M_1M_2—D$的路由。因为后者路由跳数少，源节点选择了$S—M_1M_2—D$作为发送路由。

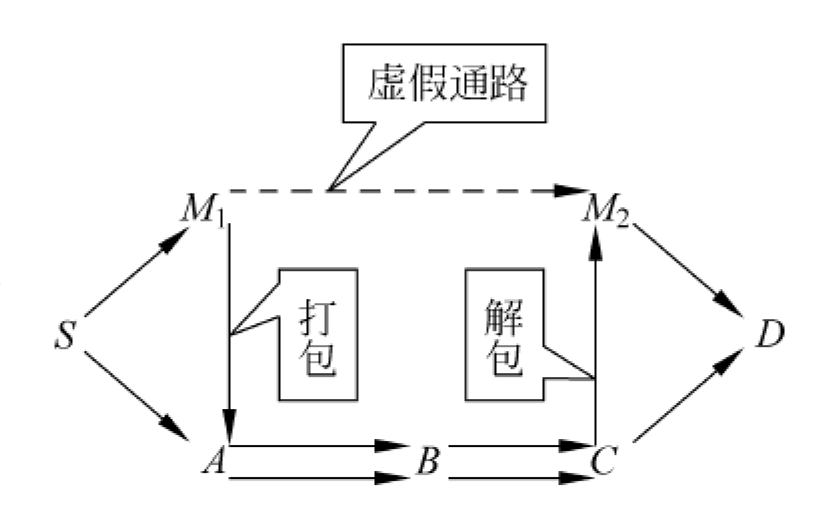


图2-2　Wormhole攻击示意图

Blackhole攻击[24]，是在路由查询中攻击者在没有至目的节点的路由情况下，抢先宣布有到目的节点的路由，使源节点建立通过该节点的路径，在随后的报文发送中，抛弃通过该节点的报文，形成抛弃报文的黑洞。

Rushing攻击[25]，在按需路由协议中，攻击者短时间内发送大量路由查询遍布全网络，使得其他节点正常的路由查询无法提交处理而被抛弃。

2.4.2 路由安全协议

1. SRP[4]

路由安全协议(Secure Routing Protocol，SRP)是对现有的按需路由协议进行扩充，实现辨别和抛弃假的路由信息，从而防止攻击者对路由信息的篡改、重放和伪造，确保获取正确的拓扑信息。协议前提为源节点与目标节点存在共享密钥，以进行认证和通信。

SRP在路由报文中扩充一个安全报头，其包含标识、序列号和报文鉴别码。当源节点发起路由请求，即路由查询报文(RREQ)时，它将源地址、目的节点地址和报文标识通过共享密钥计算出报文鉴别码，并随报文一起发送。中间节点转发报文，同时记录节点的路由请求频率，它用频率的倒数作为处理的优先级，这样可防止攻击者发出大量无用的路由请求来阻塞网络，因为这些请求的优先级将迅速降低，以至于不再处理。中间节点一般不能回答路由请求，只有当中间节点与源节点有共享密钥并有至目标的路由时才能回答路由请求。当路由请求报文到达目标节点时，目标节点首先使用共享密钥计算报文鉴别码来检验报文的完整性。如果路由请求是合法的，它会像源节点一样发出一个带有报文鉴别码的路由回答(RREP)。如果校验不通过，路由请求报文就会被扔掉。当路由答报文返回到源节点时，同样检验完整性，符合时接受其路由。路由回答报文(RRER)不需安全报头，由发现链路中断的节点直接发到源节点。

此方法的优势有以下三点。

(1) 协议简单，无须修改原路由协议，只需进行扩充就可实现安全保障。

(2) 密钥管理简单，它只需收发两端拥有密钥进行校验，网络中的节点既不拥有密钥也不参与校验，这既简化了密钥管理又减轻了节点的运算量。

(3) 适用面广,采用端到端的鉴别,可适用于多种网络协议。

它有两处不足。

(1) 路由协议为提高查找效率,中间节点能根据缓存来回答路由请求,此时因为没有与中间节点的共享密钥而无法认证。

(2) 无法认证路由维护信息,路由失败时,由中间节点产生的路由错误信息无法进行鉴别,因为源节点与路由中间节点没有共享密钥。

2. Ariadne[3]

Ariadne(A Secure On-Demand Routing Protocol for Ad hoc Networks)是一种基于DSR[24],使用 TESLA[30] 技术的安全路由协议。TESLA 是一种广播认证技术,它通过 MAC 来实现对报文的认证,利用时钟同步和密钥延迟发布来防止伪造报文鉴别码。它的基本过程是发送方先发送报文和报文鉴别码,随后再发送用于验证报文鉴别码的密钥,接收方接收后先存储报文再接收密钥进行认证。为了保证先接收报文后接收密钥的顺序,要求收发双方时钟同步。Ariadne 的前提是收发双方建立共享密钥、网络中各节点拥有其他节点的 TESLA 认证初始值,各节点时钟要求基本同步。

路由查询中实现了收方能够认证发方身份,发方能够认证路由回答报文中的每个节点,路由序列不能被篡改。收发双方通过共享密钥来实现相互认证。发方使用与收发共享密钥计算出路由查询报文的报文鉴别码,随报文一起发送。中间节点收到路由查询报文时,将本节点的标识加入到节点序列中,并重新计算节点序列的 Hash 值,然后用本节点的 TESLA 密钥计算整个报文的报文鉴别码,附在报文中转发给邻居节点。目标节点收到路由查询报文时,首先使用共享密钥认证发方身份,然后使用节点列表重新计算节点序列的 Hash 值,校验节点序列的完整性。校验通过后,目标节点形成的报文路由应答沿原路返回到源节点。路由回答报文返回过程中,各节点将自己的 TESLA 认证密钥附上。返回到源节点后,源节点通过这些密钥来校验报文的完整性。如果校验通过,则接受报文。图 2-3 显示一次路由查询的过程,S 为源节点,D 为目标节点,A、B、C 为中间节点,id 为报文标识,ti 为时间,H 代表 Hash 运算,K_A 代表节点 A 的 TESLA 认证密钥,MAC_K 代表使用密钥 K 计算报文鉴别码。底划线的部分表示与前次报文不同的部分。

路由维护时,中间节点发出 RRER 报文,并附加上本节点的 TESLA 认证信息,所有 RRER 途径节点存储该报文,等待密钥公布并认证,如果通过认证则修改路由表,否则抛弃该报文。

本方法的优势在于使用对称加密技术和广播认证技术,与非对称加密技术相比,它大大降低了节点的运算量,节省了节点资源。缺点有三点,其一要求各节点进行时钟同步,这种要求对于无线自组织网络是不现实的。其二,网络中的各节点都要发送自己的 TESLA 密钥以供其他节点认证,这需要占用不少带宽。其三,节点收到报文不能立刻认证,需要等待 TESLA 密钥的公布,这造成了一定时间的延迟。

3. ARAN[31]

ARAN(Authenticated Routing for Ad hoc Networks)适用于按需路由协议,利用公钥证书和公认的 CA 来实现认证的路由。ARAN 前提条件要求有信任的证书服务器来发放和管理证书,每个节点加入网络前必须从证书服务器获取一个公开密钥的证书。

S: $h_0 = MAC_{K_{SD}}(REQUEST, S, D, id, ti)$

S−>*: $(REQUST, S, D, id, ti, h_0, (), ())$

A: $h_1 = H[A, h_0]$ $M_A = MAC_{K_{A_{ti}}}(REQUEST, S, D, id, ti, h_1, (A), ())$

A−>*: $(REQUEST, S, D, id, ti, \underline{h_1}, (\underline{A}), (\underline{M_A}))$

B: $h_2 = H[B, h_1] M_B = MAC_{K_{B_{ti}}}(REQUEST, S, D, id, ti, h_2, (A, B), (M_A))$

B−>*: $(REQUEST, S, D, id, ti, \underline{h_2}, (A, \underline{B}), (M_A, \underline{M_B}))$

C: $h_3 = H[C, h_2]$

$M_C = MAC_{K_{C_{ti}}}(REQUEST, S, D, id, ti, h_3, (A, B, C), (M_A, M_B))$

C−>*: $(REQUST, S, D, id, ti, \underline{h_3}, (A, B, \underline{C}), (M_A, M_B, \underline{M_C}))$

D: $M_D = MAC_{K_{DS}}(REPLY, D, S, ti, (A, B, C), (M_A, M_B, M_C))$

D−>C: $(REPLY, D, S, ti, (A, B, C), (M_A, M_B, M_C), \underline{M_D}, ())$

C−>B: $(REPLY, D, S, ti, (A, B, C), (M_A, M_B, M_C), M_D, (\underline{K_{C_{ti}}}))$

B−>A: $(REPLY, D, S, ti, (A, B, C), (M_A, M_B, M_C), M_D, (K_{C_{ti}}, \underline{K_{B_{ti}}}))$

A−>S: $(REPLY, D, S, ti, (A, B, C), (M_A, M_B, M_C), M_D, (K_{C_{ti}}, K_{B_{ti}}, \underline{K_{A_{ti}}}))$

图 2-3 Ariadne 协议的路由查询过程

路由查询时，源节点发出一个经过签名的路由查询报文。第一跳节点校验源节点签名后，签名并附上自己的证书。随后每个转发节点都先校验前一个节点的签名，然后用自己的签名和证书替代上个节点的签名和证书。路由查询报文到达目标后，目标节点校验签名，然后产生路由应答，沿来路返回，途径的每个节点与来时一样进行校验和签名。源节点收到路由应答，校验目的节点的签名正确后接受其路由。路由中断时，路由维护报文由发送节点签名后直接转发至源节点，中间节点不再进行修改。图 2-4 显示 ARAN 路由查询中如何进行签名和证书的更替，S 为源节点，D 为目标节点，A、B 为中间节点，$CERT_S$ 为源节点的证书，N 为一个随机整数，t 为发送时间，N、t 用于防止重放攻击。

S−>*: $[REQUEST, D, CERT_S, N, t]K_{S-}$

A−>*: $[[REQUEST, D, CERT_S, N, t]K_{S-}]K_{A-}, CERT_A$

B−>*: $[[REQUEST, D, CERT_S, N, t]K_{S-}]K_{B-}, CERT_B$

图 2-4 ARAN 路由查询

本协议的优点有两点，其一，ARAN 使用公开密钥算法实现了报文鉴别、完整性和不可抵赖性。在路由查找和路由应答中使用端到端和单跳的认证来阻止伪造，使用数字签名来阻止报文篡改。从安全的角度来看，它提供了较完善的安全功能。其二，路由报文经过的每一个节点都相互签名认证，所以攻击者没有机会加入网络进行攻击。缺点有四点，其一为中间节点不能回答路由请求，必须由目的节点来回答，降低了路由协议的效率。其二为使用公开密钥算法来进行认证，会导致节点计算负载过重，如果攻击者发出大量路由请求，会导致节点来不及进行认证而无法处理正常的路由信息。其三，如果被攻破的节点从内部发动攻击，该算法不能抵抗，因为攻击者拥有合法的证书。其四，要求有一个公共的 CA 来维护每一个节点的证书，该 CA 失败将导致整个网络的安全失效。

4．SEAD[32]

SEAD(Secure Efficient Distance Vector Routing for Mobile Wireless Ad hoc Networks)是基于距离向量 DSDV[26] 的安全路由协议，基本思想是利用 Hash 链中的元素来认证路由更新报文中的序列号和跳数。所谓单向 Hash 链，即先选择一种 Hash 函数，如 MD5[33]，再选择一个 0～1 之间的随机数 X，令 $H_0=X$，$H_1=\text{MD5}(H_0)$，$H_2=\text{MD5}(H_1)$，依此计算出一系列数 $H_0,H_1,H_2,\cdots,H_n$，这一系列数就叫单向 Hash 链。利用这一系列数就可实现认证，如已知 H_8，要鉴别 H_5 是否合法，只要计算出 $\text{MD5}(\text{MD5}(\text{MD5}(H_5)))$ 与 H_8 相对照即可。因此，只要拥有认证元素 H_n，所有 Hash 链中的元素均可鉴别。

SEAD 的前提是有信任的实体来分配和维护各节点的认证元素。每一个节点在路由更新中从自己的 Hash 链中选出一个元素用于认证。节点选择 Hash 值的算法为，假定网络最大跳数为 m，Hash 链为 $H_0,H_1,H_2,\cdots,H_n$，将 Hash 链分为 n/m 组，每一组用于认证一个序列号，组内的 Hash 值用于认证跳数。例如：序列号 i，令 $k=n/m-i$，其组内元素为 H_{km}，$H_{km+1},\cdots,H_{km+m-1}$，如果跳数为 j，则 Hash 值 H_{km+j} 即可用于认证序列号为 i 跳数为 j 的路由更新。由于单向 Hash 链的特性，能够防止攻击者伪造一个比真实序列号大的报文，或比真实跳数小的报文。当节点收到路由更新报文时，它首先利用 Hash 值进行认证，如果认证通过则修改路由表，否则抛弃该报文。此方法的优点为采用单向 Hash 链算法进行认证，大大降低了节点的运算复杂度。缺点为在网络运行整个过程都需要有信任的实体来分配和维护每个节点的认证元素。因为 Hash 链中的元素会被用完，用完后节点要重新计算一条 Hash 链，其认证元素 H_n 要由信任实体重新分配给所有节点。该信任实体容易引起单点失败，它若被攻击者攻破了，则整个网络路由协议无法认证了。

5．SAODV[34,35]

SAODV(Secure Ad hoc On-Demand Distance Vector)是基于 AODV[25] 的路由安全协议，它的前提条件是要将网络中所有节点的公钥分发到各节点，以便用于签名认证，它使用两种机制来保证 AODV 的协议安全。一种是数字签名，用来保证报文中不需变化部分的完整性，提供端到端的鉴别。另一种是单向 Hash 链，用来保证路由报文中可变的部分如跳数的认证。

源节点路由查找时，发出带有数字签名和 Hash 值的路由查询报文，中间节点收到路由查询报文时，首先校验数字签名和 Hash 值，能通过时处理该报文，否则抛弃该报文。路由查询报文到达目的节点时，节点形成路由回答报文，同样进行数字签名和计算 Hash 值后沿来路返回。与 ARAN 不同的是中间节点并不需要相互签名认证，只有源目的节点对路由报文进行数字签名，中间节点只需验证数字签名，而不需形成数字签名，计算量比 ARAN 少了一半。

对于 AODV 路由查找中，中间节点如何回答路由查询的问题，采用双签名方法来解决，即源节点在发出路由查询报文中设立一个标志位并带上一个返回路由回答报文的签名，这样，中间节点可根据附带的签名生成报文返回源节点。对于路由出错报文的处理方法为，产生该报文的节点进行数字签名，这样就可防止攻击者编造报文。

该协议的优点在于采用双签名机制解决了中间节点回答路由请求的问题。缺点是采用公开密钥算法，中间转发节点需要检验数字签名，计算负载比较大。

6. SLSP[36]

SLSP(Secure Link State Routing for Mobile Ad hoc Networks)是基于链路状态的路由安全协议。它保护使用链路状态算法的路由协议,例如:ZRP[37]。算法前提为每个节点持有公钥和私钥,并将公钥发送给所有节点。SLSP 对链路状态更新报文扩充一个安全报头,通过数字签名来提供认证和完整性,报文序列号来防止重放攻击,单向 Hash 链来限制转发次数。各节点周期性地向网络节点广播经过签名的链路状态更新报文。网络中节点收到链路状态更新报文后,首先检查签名和报文完整性,如果检查通过则接受该报文,若没达到最大转发次数则转发该报文,如果检查没通过则抛弃该报文。SLSP 还包括邻居监视机制,每个节点将其 MAC 地址和 IP 地址经过签名后发给邻居节点,邻居节点记录相应地址。它有以下两个用途。

(1) 它可防止伪造 IP 地址。

(2) 用来记录邻居发送报文的频率,如果发送报文的频率过高,超过一定限额,就可以认定为攻击者,对其发出的报文不再处理而直接抛弃,这样就可将攻击者滥发的报文限制在单跳邻居范围之内,有效地防止了拒绝服务攻击。

该协议的优点在于采用邻居监视机制来防止拒绝服务攻击。缺点是采用公开密钥算法,各节点既要生成本节点报文的数字签名又要检验其他节点报文的数字签名,计算负载比较大。

2.4.3 路由安全协议的比较与分析

表 2-2 对本文介绍的六种比较典型的路由安全协议进行了分析和比较。它们有一些共同的特点,需要事先通过信任的实体或证书服务器将密钥、证书或认证元素发布到各节点,协议都能够抗击前述的各种网络进攻,如冒充、伪造等,但都只局限于攻击者的单个进攻,针对联合进攻,如 Wormhole[29] 攻击,各个协议都无法对付,文献[38]专门提出一种方法 Packet leashes 抵抗 Wormhole 攻击,它基于精确的时间或位置信息来发现并阻止它的攻击。

表 2-2 路由安全协议的比较

协议名称	适用的路由协议	协议前提	主要的安全技术	认证部分	优 势	缺 点
SRP	DSR	源节点与目标节点建立共享密钥	报文鉴别码	源地址、目的地址、报文标识	算法简单、适用面广	缺乏对路由维护信息的保护、中间节点不能回答路由请求
ARIADNE	DSR	发布 TESLA 认证密钥、源目标节点建立共享密钥、各节点时钟同步	单向 Hash 链、报文鉴别码	整个报文、路由序列	采用对称密钥和 TESLA 技术,计算量小、管理简单	要求节点时钟同步、发送认证密钥占用带宽、认证有延迟

续表

协议名称	适用的路由协议	协议前提	主要的安全技术	认证部分	优　势	缺　点
ARAN	AODV DSR	建立证书服务器,发布和维护每个节点的公钥证书	数字签名	整个报文	实现了鉴别、完整性和不可否认性	计算量大、需要信任的CA、中间节点不能回答路由请求
SEAD	DSDV	发布认证初始值	单向Hash链	序列号、跳数	计算负载小	需要信任的实体来分配和维护各节点的认证元素
SAODV	AODV	分发节点公钥	数字签名、单向Hash链	整个报文	中间节点可以回答路由请求	采用公开密钥算法计算量大
SLSP	ZRP	分发节点公钥	数字签名、单向Hash链	整个报文	采用邻居监视机制来防止拒绝服务攻击	采用公开密钥算法计算量大

以上的路由安全协议存在的问题及改进方法:

(1) 有些协议在设计上强调了安全性,而忽视了可用性。没有充分考虑到无线自组织网络节点计算能力弱、电池和通信带宽有限的特点,如ARAN、SAODV、SLSP协议,采用公开密钥证书加数字签名的安全机制,在安全方面是完善的,但由于数字签名的生成和检验都是计算量非常大的、非常耗时的,这对于本身计算能力弱,同时还要承担转发报文和运行应用程序的移动节点来说,是一项沉重的负担,如果大量报文同时到来,节点就会因为检验数字签名过慢而来不及处理报文,导致拒绝服务。安全协议的设计上要充分考虑到无线自组织网络资源有限的特点,尽量采用报文鉴别码和单向Hash链等运算量小的安全技术,以减少节点的运算时间和能量消耗。不宜采用数字签名等计算量大的算法。

(2) 设计上为了保证安全性,屏蔽了路由协议的某些功能,降低了路由协议的有效性。例如SRP、ARAN协议为保证安全、降低算法的复杂性都取消了中间节点回答路由请求的功能,必须由目标节点产生路由回答,降低了路由查询的效率。为了解决该问题,可采用TESLA认证或共享密钥来实现对中间节点路由回答认证。

(3) 有些协议的前提条件要求在网络运行过程中需要集中的服务器支持。如ARAN要求有证书服务器管理各节点公钥证书,SEAD要求信任的实体来分配认证元素。集中的服务器虽然保证了协议的安全,但却带来了单点失败的威胁,如果该服务器被攻破,则整个网络安全就失效了。因此,在设计上应尽量不用集中的服务器,如果要用也只能在初始阶段,网络运行时或者不用或者用分布式的CA来实现。

(4) 有些协议提出的要求是网络难以提供的。如Ariadne协议要求网络所有节点时钟同步,这对于大型网络是难以实现的。

2.5　入侵检测

无线信道、动态拓扑、合作的路由算法、缺乏集中的监控等都使得无线自组织网络安全更加脆弱,特别是移动节点缺乏物理保护,容易被偷窃、捕获,落入敌手后重新加入网络,导

致攻击从内部产生。而采用密码学理论的网络安全方案无法对抗此类攻击。目前还没有100%的安全方案,无论多么安全的方案都可能存在这样或那样的漏洞。因此,入侵检测就理应成为安全方案之后的第二道防护墙。

2.5.1 入侵检测方案

Yongguang Zhang 和 Weeke Lee 提出了一个基于代理的分布式协作入侵检测方案[39]。在该方案中 IDS 代理运行于网络中每一个节点上,拥有六大功能模块,分为数据收集、本地检测、合作检测、本地入侵响应、全局入侵响应、安全通信。图 2-5 为 IDS 代理由六大功能模块组成的示意图。其过程为首先执行本地数据收集和检测。如果本地节点能够确定入侵已发生,则直接告警。如果只是怀疑有入侵行为,本地节点能够激发多节点的协作检测,进一步是否发生了入侵。如果确定有入侵则激发全网的入侵响应。同时提出了一个检测路由进攻的异常检测模型,通过提取正常网络运行时的数据,进行分类训练,实现对路由入侵的检测。为了提高检测效率,入侵检测并不局限于网络层,而是多层综合检测。Yongguang Zhang 在后续文献[40]中对上述方案进行了详细的论述,建立了一个 IDS 模型并用网络模拟器实现了模拟运行。

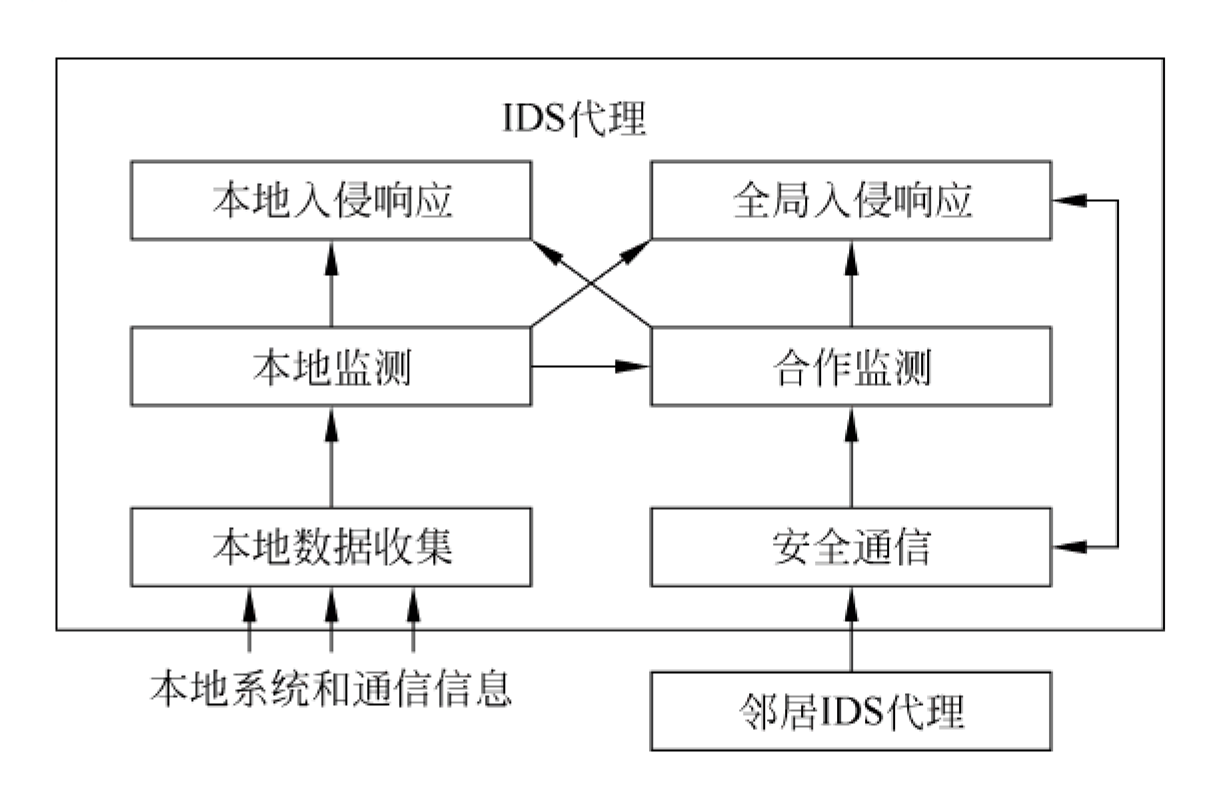


图 2-5 IDS 代理组成

上述方案的优势有两点,其一,提出了分布式协作入侵检测的架构,利用分布在每个节点的 IDS 代理独立完成本地检测,合作完成全局检测,适合于无线自组织网络自组织的特点。其二,采用多层综合入侵检测,提高了检测效率。缺点也有两点,其一,采用异常检测模式,要事先采样数据进行训练,不适合于无线自组织网络多变的应用场合。其二,每个节点都运行有代理,占用过多的内存和计算资源。

Oleg Kachirski 和 Ratan Guha 提出了基于移动代理的入侵检测方案[41]。他们认为 Yongguan Zhang 的方案每个节点都有代理,过于占用网络资源,为了节省资源,只是在某些节点上驻留有监视网络的代理,并且代理的数量可按要求进行增减。

Chin-Yang Tseng 等有提出了基于规范(specification-based)入侵检测方案[42]。该方案利用分布在网络中的监测点,合作监视在 AODV 路由查询过程中,被监视节点是否按路由规范进行操作。如果发现不一致则报警。检测过程为,监听节点对查询报文的处理过程,记录下来形成转发表和操作树,然后用规范形成的有限状态机进行检查,输出为正常状态、怀

疑状态、入侵状态三种结果，再分别进行不同的处理。该方案优点在于采用了基于规范入侵检测，既不需要事先提取入侵行为特征，也不需要数据进行训练，有较高的检测率和较低的误报率。缺点为，占用节点较多的计算资源，也未用实验进行验证。

2.5.2　入侵检测方案的比较与分析

表 2-3 对三种入侵检测方案进行了比较，上述入侵检测方案存在以下一些问题：

表 2-3　三种入侵检测方案的比较

协议名称	基于代理的分布式协作入侵检测	基于移动代理的入侵检测	基于规范入侵检测
执行者	驻留节点上的代理	各种移动代理	每个节点
检测模式	异常检测	异常检测	基于规范的检测
检测方法	分布式监测、邻居监视	分布式监测、邻居监视	分布式监测、邻居监视
优点	各代理合作监测与响应	可动态调整代理数量，降低对资源的消耗	不需要数据进行训练，较高的检测率
缺点	占用过多资源	协议比较复杂	计算量大

(1) 现行的入侵检测的架构为使用代理作为入侵检测的执行者，代理驻留并运行于网络中每一个节点内，分布式的监视网络状况，信息共享，合作检测入侵行为。这种架构对于入侵检测来说是较为有效的，但未充分考虑到网络带宽和节点计算资源有限的特点，节点本身要运行自己的应用程序，又要负责网络报文转发，CPU 和内存资源已经很紧张，还要运行代理监视网络和主机，这也许会导致节点性能明显下降，甚至资源枯竭。网络带宽也十分有限，代理间信息交换也要占用带宽，这也许会影响网络性能。我们认为在设计上应充分考虑到网络资源有限的特点，降低其对资源的要求，不必每一个节点都运行代理，可采用两种方式，一种是分区域，每个区域使用一个代理负责监控。另一种是使用少量移动代理散布于网络各处，如发现异常，可向异常处移动，进一步检测以确定是网络故障还是入侵行为。

(2) 入侵检测的模式通常为异常检测模式。异常检测模式是定义好正常行为的范围，凡是偏离了正常的行为都为入侵行为。该检测模式因为其能够检测出新的入侵行为而被采用。但在无线自组织网络环境下，因为动态的拓扑、无线信道的不稳定、应用环境的多变，使得难以准确定义正常行为的范围。如果定义不当，将会导致不能发现入侵行为或者错误报警率太高。针对此不足，应该采用基于规范的入侵检测模式，事先按网络协议的规范定义好程序的执行步骤，运行时监视程序是否按规范执行，偏离了即为入侵行为。无线自组织网络协议都有明确的规范，使用该技术既能检测出未知的入侵行为，又有较低的误报警率。

2.6　增强合作的机制

无线自组织网络不像固定网络，有专门的路由器和交换机来实现网络通信功能，它的每一个节点既是主机又是路由器，既作为一个网络终端用户又作为一个网络交换节点，因此，每一个节点要承担起网络路由和包交换的功能。但在无线自组织网络中每个节点拥有的资源有限，在多个管理域的情况下，有些节点为了节省自己的资源，不参与网络交换，这就是无

线自组织网络节点中的自私行为。这种自私行为对网络性能的影响不可低估。在文献[43]中研究了在DSR协议环境下,自私行为对整个网络吞吐量和传输延迟的影响程度,模拟试验结果显示,既使整个网络节点中只有小部分节点产生自私行为,也造成网络性能的严重下降。自私行为和攻击者的蓄意破坏行为虽然在出发点上有所不同,前者是为了保存和节省资源,后者是为了破坏网络的正常功能,但它们所造成后果相同,都会严重影响网络性能,所以如何对付自私行为,增强网络节点的合作机制也应该是网络安全的一个研究方向。在对付自私行为,增强合作方面的论文主要分为两类,第一类是基于激励的机制,其基本原理为节点转发报文后,即可得筹码或虚拟货币用于自己报文的发送。第二类是基于惩罚的机制。邻居相互监视,发现不良行为的节点,则将被排除出网络。下面首先介绍这两类算法,然后进行分析比较。

2.6.1 基于激励的机制

Levente Buttyan 和 Jean-Pierre Hubaux 在文献[44]中,提出了为了增强合作,网络中的节点必须被鼓励转发报文,但又不能滥发报文增加网络负载。为此,设计两种虚拟货币的解决方案。一种是钱包方式,需要发送报文的节点估计报文所经过的节点与所需要的花费,将计算所得的虚拟货币数放入钱包,钱包随报文一起发送,途径中间的每一个转发节点从报文的钱包中取出一定量的货币作为转发该报文的费用,然后转发该报文,直至到达目的节点。这样每一个节点只有尽力转发其他节点的报文才能获得足够的货币以发送自己的报文,从而起到一种激励合作的作用。该方式的优点是可以防止节点滥发报文增加网络负载,缺点有两点,其一是源节点需要能够精确计算报文转发的费用,如果放入钱包的费用小于实际费用,报文就会被中途抛弃,在拓扑频繁变化的无线自组织网络中估计报文所经过的节点和费用不是一件容易的事情。其二是每个报文都要携带货币,用于中途付给转发节点,增加了报文的长度。另一种解决方案是购买方式,每一个节点转发报文时,从上游节点买下该报文,加上本节点的转发费用后,又把报文卖给下游节点,依此转发直至目标节点。该方式的优点在于无需事先估计路途转发费用,全部费用由收方支付,缺点为由于发方无需支付发送费用,攻击者可滥发报文,造成网络负载过重直至崩溃。

Levente Buttyan and Jean-Pierre Hubaux 在文献[27]中,提出了一种基于筹码的方案。该方案要求每个节点都安装一个防止用户修改的硬件叫安全卡。该卡内有筹码累加器,每当节点转发一个报文时,该计数器就增加一个筹码值。当节点需要发送自己的报文时,就要将卡内筹码累加器减去 n 个筹码,n 代表报文要经过 n 个节点转发才能到达目的节点。如果卡内筹码数小于 n,则节点不能发送该报文。这意味着节点要发送自己的报文,首先必须转发其他报文,以积累足够筹码才行。该方案并不是用于阻止节点的不良行为,它只是鼓励节点转发报文,确保节点不能从不良行为中获益,如果节点不转发报文,它就没有筹码用于发送自己的报文。其优点是算法简单,只需要一个筹码累加器即可。缺点有两点,其一需要硬件支持。其二不论报文长短,转发时均增加一个筹码。也许转发一个长报文的花费相当于转发几个短报文的花费。

Sheng Zhong 等人认为在每个节点安装额外硬件是不现实的,他们提出一个不需要在节点安装硬件的类似方案 sprite[45]。首先设立一个集中的结算中心来存储每个节点的筹码数,当节点发送一个报文时,所有中间转发节点和目的节点都记下这个报文的收据,然后与

结算中心联系，上传该收据，结算中心根据报文转发情况付给参与转发节点相应的筹码数，同时扣去发送节点总计筹码数。结算时，按照博弈理论来计算支付方案，促使节点能够尽力诚实履行网络功能。该方案的优点在于用统一的结算中心取代了各节点的硬件累加器，无需在每个节点安装硬件卡。缺点也在集中的结算中心，若有网内节点承担则会造成过重的通信负载和单点失败。该算法提出使用网外设备，将结算中心放在网外，通过移动通信中的GPRS来实现与结算中心的联系。这既增加了硬件设备，又限制了网络的应用场合。

该类方案有三个特点，其一使用虚拟货币或筹码作为转发的回报。其二使用硬件设备来存储筹码值，以防用户修改。其三采用博弈理论来促使节点合作。

2.6.2　基于惩罚的机制

Sergio Marti 等人提出通过 watchdog 和 pathrater 两种技术来对付不良节点的方案[46]。Watchdog 和 pathrater 运行于每一个节点上，Watchdog 负责监视邻居节点行为，发现不良行为的节点。Pathrater 负责选择避开不良节点的路径。该方案只是实现如何发现并避开不良节点，并不孤立和惩罚不良节点。不良节点仍然能够正常收发报文，反而为其免除了正常的转发流量，客观上奖赏了不良节点。

Sonja Buchegger 和 Jean-Yves Le Boudec 在文献[47，48]中提出一种利用邻居监视来发现并排除不良节点的算法—CONFIDANT。它依靠运行于每个节点上的四个程序来实现其功能。邻居监视器：用于发现不履行正常网络功能的邻居节点。信任管理：用于发送、接收、管理其他节点的报警信息。名誉系统：标记并管理其他节点的名誉值。路径管理：路径选择时避开并孤立不良节点。每个节点监视其邻居的行为，如果发现不良行为则提交给名誉系统并给发送报警信息给其他节点。名誉系统为每个节点设立一个名誉值，当有不良行为报告时，减少其名誉分值。当某个节点的名誉值低于标准时，则提交路径管理。路径管理将删除与不良节点有关的路径存储信息，并拒绝其路由申请。该方案的优点是不仅发现不良节点，而且用孤立方法来惩罚它们，以促使它们履行正常网络功能。

Pietro Michiardi 和 Refik Molva 也提出一种通过监视技术和名誉机制来激励合作的算法 CORE[49]。网络中每一个节点监视其邻居的行为，观察其是否履行正常网络功能，如转发报文、处理路由请求等。如果观察结果与预期的一致，则增加该节点的名誉分值，否则减去一定分值。节点通过一个综合公式来计算某个节点总名誉分值，公式的参数有直接观察结果、其他节点的报告等，还考虑到通信线路状况和以前的名誉分值。当某个节点出现自私行为或其他恶意行为时，其名誉分值会逐渐下降，当低于某个值时，其他节点会拒绝为其提供服务，那样就会将自私节点排除出网络。

这类方案有三个特点，其一采用本地监视技术来发现不良节点。其二通过名誉值来评价节点。其三通过惩罚不良节点来促使节点合作。

2.6.3　两类算法的比较与分析

为了实现共同的目标，限制节点的自私行为，增强节点合作，提高网络性能，上述两类算法采用完全不同的方法，一种是激励的方法，另一种是惩罚的方法，表 2-4 对两类算法进行了比较。

表 2-4　基于激励的机制算法与基于惩罚的机制算法的比较

算法分类	基于激励的机制	基于惩罚的机制
理论基础	博弈论	
实现手段	虚拟货币	邻居监视
主要思想	以虚拟货币作为节点合作的奖赏，鼓励转发报文	以被排出网络作为不合作的惩罚，鞭策节点参与网络功能
基本原理	节点转发报文后，即可得筹码或虚拟货币用于自己报文的发送	邻居相互监视，发现不良行为的节点，则将被排除出网络
优点	算法简单	纯软件实现，无需硬件支持
缺点	需要硬件支持	邻居监视并不完全有效

上述算法存在问题及改进方法如下：

（1）基于激励的机制算法存在的主要问题是需要额外的硬件支持其协议的执行。前两种方案需要在每一个节点安装安全卡用来存储信息，以防用户修改，后一种方案需要使用网络之外的设备来统一存储每个节点的筹码信息。这些设备的使用会限制网络的扩展性和应用范围。例如多个管理域的网络中用户若不同意安装安全卡，就无法运行前两种协议。无线自组织网络布置于没有移动通信支持的场合，后一种方案就无法实施。其解决方法应该采用纯软件的方法来实现整个算法，对于筹码信息的管理可设计专门的代理运行于每个节点内，代理独立于用户运行，设立安全措施防止用户更改数据。

（2）基于惩罚的机制算法中存在邻居节点监视的有效性问题。基于惩罚的机制主要使用本地邻居监视的方法来发现不良行为，但无线自组织网络具有多变的拓扑和不稳定的无线通信的特征，这使得有效检测自私节点变得十分困难。节点可能由于线路和电源等故障而无法转发报文，被认定为自私节点。解决这种问题可采用多种参数综合评价的方法，例如，将一个节点的名誉值与其过去的行为相关，如果一个节点过去行为一直良好，近期突然不转发报文了，并不立即认定其为不良节点，也许是线路故障，需要等待一段时间再作评价。

（3）标识的有效性问题。无线自组织网络中节点可不断变动其位置。当在某处被邻居认定为自私的节点而被排除出网络，它可移动到新的地点，通过改变节点标识，又可重新加入网络。解决这种问题可采用两种方式，一种是利用用户公钥来生成用户标识，另一种是直接利用标识作为用户的公钥。这两种方式都能防止用户随意改变其标识。

2.7　小结

由于无线自组织网络的独特结构，使得常规的安全方案无法应用，必须针对其特点设计专门的安全解决方案。本文从密钥管理、路由安全、入侵检测、增强合作几个方面介绍了应用于无线自组织网络的安全解决方案。首先讨论了密钥管理，主要介绍了自组织的密钥管理和分布式的密钥管理两类算法，指出了其优点和缺点。然后分析了五种典型的路由安全协议，对它们进行了综合比较并指出其存在问题及改进方法。接下来说明了基于代理的分布式监视合作检测的入侵检测体系结构。最后讨论了基于激励和基于惩罚的两种增强合作的机制。

无线自组织网络安全的研究是一个年轻而又迅速发展的领域，总体来说，下一步发展应

包括以下几个方面：

(1) 进一步提高性能，降低算法对资源的要求。无线自组织网络中节点本身的计算能力和电池能量都十分有限，还要参与网络交换。网络安全作为网络正常运行的一种保障，不应该也不允许占用节点大量的资源，不能因为增加了安全措施，降低了网络性能，影响了网络的正常运行。应该设计和采用一些对资源要求少的算法，如：使用本地认证取代分布式的认证，用对称密钥取代公开密钥等。

(2) 增强协议的可扩展性。有些协议在节点数目较少时，性能还可以，当节点数目增加时，其性能会明显下降。如：自组织的密钥管理当网络扩大时，证书数据库的形成、维护和认证的花费会明显增加。算法在设计时，就应考虑到其可扩展性。

(3) 安全方案应具有自适应可调整的特性。安全方案不应该是固定的，它应该具有自适应的性能，根据网络的资源状况调整其功能，资源充足时，功能强一些，反之则功能减少一些，也可以将一些占有资源较多的功能模块分布到一些资源多、性能好的节点上去，使得整个网络负载均衡。Jiejun Kong 等人在这方面进行了一些研究，提出了一种自适应的安全方案[50]。平时，利用无人飞机上的节点来实现集中的 CA 或 KDC，因为飞机上的节点资源相对充足。战时，当无人飞机被摧毁时，集中的 CA 或 KDC 自动转化为分布式的，由 n 个移动节点承担。虽然性能会有所降低，但仍然能保证整个网络的安全运行。

(4) 提供对多播的安全保护。多播的应用能够有效地减少网络流量，特别适用于军事指挥网络。现行大多数的安全方案只停留在如何保护路由信息的完整性，如何实现对单个节点的认证，没有考虑如何实现对多播的安全支持。只有少数文献[51,52]论述了安全的多播，对无线自组织网络环境下的安全多播的研究还很不完善，需要进一步发展。

(5) 在无线自组织网络安全的研究领域内，除了本文论述的五个主要的研究方面，还有许多新领域有待于去拓展：

① 如何保护节点通信量和位置的信息。通过通信量的分析能够确定网络中节点的角色，再确定节点的位置，就可将攻击指向网络的要害，如网控中心、集中的 CA 或军事指挥网中指挥员等。

② 各种针对中路由协议的攻击及对策。如：wormhole、rushing 攻击等，因为无线自组织网络的复杂性，也许还存在许多新类型的攻击尚未发现。

③ 路由安全算法研究主要集中在 DSR、AODV、DSDV 路由协议上，还应拓展其范围，设计一些其他路由安全算法，如基于位置的路由安全协议、基于能量的路由安全协议、层次路由安全协议等。

④ 链路层和高层的安全协议的研究。

⑤ 如何实现网络的存取控制。

参考文献

[1] Macker C J. Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations, RFC 2501, January 1999.

[2] Buttyan L, Hubaux J P. Report on a Working Session on Security in Wireless Ad hoc Networks. Mobile Computing and Communications Review, 2002, 6(4).

[3] Hu Y C, Adrian Perrig, Johnson D B. Ariadne: A secure On-Demand Routing Protocol for Ad hoc Networks, in Proceedings of the MobiCom, Atlanta, Georeia, 2002, September: 23-28.
[4] Papadimitratos P, Haas Z, Secure Routing for mobile Ad hoc Networks, in Proceedings of the SCS communication Networks and Distributed Systems Modeling and Simulation Conference, San Antonio, TX, 2002: 27-31.
[5] Jean-Pierre Hubaux, Levente Buttyan, Srdjan Capkun, The Quest for Security in Mobile Ad hoc Networks, In Proceedings of the 2001 ACM International Symposium on Mobile Ad hoc Networking & computing 2001, Long Beach, CA, USA.
[6] Srdjan Capkun, Levente Nuttyan, Jea-Pierre, Self-organized Public-Key Management for Mobile Ad hoc Networks, IEEE Transactions on mobile computing, 2003, 2(1).
[7] Zimmermann P. The Official PGP User's Guide, MIT Press, June 1995.
[8] Desmedt Y. Threshold cryptography, European Transactions on Telecommunications, July-August 1994, 5(4): 449-457.
[9] Lidong Zhou, Haas Z J. Securing Ad hoc Networks, IEEE Networks Special Issue on Network Security, November/December, 1999.
[10] Ostrovsky R, M Yung. How to Withstand Mobile Virus Attacks, In Proceedings of the 10th ACM Symposium on Principles of Distributed Computing, 1991: 51-59.
[11] Jiejun Kong, Petros Zerfos, Haiyun Luo, et al. Providing Robust and Ubiquitous Security Support for Mobile Ad-hoc Networks, IEEE 9th International Conference on Network Protocols (ICNP'01), 2001.
[12] Luo Haiyun, Kong Jiejun, Petros Zerfos. et al, Self-securing Ad hoc Wireless Networks, In Proceedings of the Seventh IEEE Symposium on Computers and Communications (ISCC'02).
[13] Aram Khalili, Jonathan Katz, William Arbaugh. Towards Security Solutions for Truly Ad-hoc Networks, IEEE Workshop on Security and Assurance in Ad hoc Networks, in Conjunction with the 2003 International Symposium on Applications and the Internet, Orlando, FL, January 28, 2003.
[14] Douceur J. The Sybil attack, In Proceedings of the 1st International Workshop in Peer-to-Peer Systems(IPTPS), 2002.
[15] Frank Stajano, Ross Anderson. The Resurrecting Duckling: Security Issues for Ad-hoc Wireless Networks, In Proceedings of the 7th International Workshop on Security Protocols (1999), LNCS 1796, Springer-Verlag, Berlin Germany, April 1999.
[16] Akyildiz I F, Su W, Sankarasubramaniam Y, et al. Wireless Sensor Networks: A Survey, Computer Networks, 2002, 38(4): 393-422.
[17] Asokan N, Philip Ginzboorg. Key agreement in Ad hoc Networks, Computer Communications, 2000, 23: 1627-1637.
[18] Klaus Becker, Uta Wille, Communication Complexity of group key distribution, In 5th ACM Conference on Computer and Communications Security.
[19] Zheng Yan. Security in Ad hoc Networks, http://citeseer. nj. nec. com/536945. html.
[20] Srdjan Capkun, Jean-Pierre Hubaux, Levente Buttyan. Mobility Helps Security in Ad hoc Networks. The Fourth ACM International Symposium on Mobile Ad hoc Networking and Computing Annapolis, Maryland, USA 2003: 1-3.
[21] Tuomas Aural, Silja Mäki. Towards a Survivable Security Architecture for Ad-hoc Networks, Microsoft Research, Lecture Notes in Computer Science 2002, 2467(0302: 9743).
[22] Dasgupta P, Gokhale S. Distributed Authentication for Peer to Peer Networks, IEEE Workshop on Security and Assurance in Ad hoc Networks, In Conjunction with The 2003 International Symposium on Applications and the Internet, Orlando, FL, January 28, 2003.

[23] Lakshmi Venkatraman, Agrawal D P. A Novel Authentication scheme for Ad hoc Networks, Wireless Communications and Networking Conference (WCNC 2000), IEEE, Pages 1268-1273, 3.
[24] Johnson D B, Maltz D A, Hu Y C. The Dynamic Source Routing Protocol for Mobile Ad hoc Networks (DSR), INTERNET-DRAFT, draft-ietf-manet-dsr-10. txt, 19 July 2004.
[25] Perkins C, Belding-Royer E, Das S. Ad hoc On-Demand Distance Vector (AODV) Routing, RFC3561, July 2003.
[26] Perkins C E, Bhagwat P. Highly Dynamic Destination Sequenced Distance-Vector Routing (DSDV) for mobile computers, The ACM SIGCOMM Conference on Communications Architectures, London, 1994.
[27] Buttyán L, Hubaux J P. Stimulating Cooperation in Self-Organizing Mobile Ad hoc Networks, ACM Journal for Mobile Networks(MONET), Special Issue on Mobile Ad hoc Networks, summer 2002.
[28] Sergio Marti, Giuli T J, Kevin Lai, et al. Mitigating routing misbehavior in mobile Ad hoc networks. In Proceedings of the Sixth International Conference on Mobile Computing and Networking (Mobicom), Boston, August 2000.
[29] Hu Y C, Perrig A, Johnso D B. Wormhole Detection in Wireless Ad hoc Networks, Technical Report TR01-384, Department of Computer Science, Rice University, December 2001.
[30] Perrig A, Canetti R, Song D, et al. Efficient and Secure Source authentication for multicast, in Proceedings of Network and Distributed System Security Symposium, February 2001: 35-46.
[31] Kimaya Sanzgiri, Bridget Dahill, Brian Neil Levine, et al. A Secure Routing Protocol for Ad hoc Networks, In Proceedings of 2002 IEEE International Conference on Network Protocols (ICNP), November 2002.
[32] Hu Y C, Johnson D B, Adrian Perrig. SEAD: Secure Efficient Distance Vector Routing for Mobile Wireless Ad hoc Networks, in Proceedings of the 4th IEEE Workshop on Mobile Computing Systems & Applications (WMCSA 2002), pp. 3-13, IEEE, Calicoon, NY, June 2002.
[33] Rivest R L. The MD5 Message-Digest Algorithm, RFC1321, April 1992.
[34] Manel Guerrero Zapata, Secure Ad hoc On-Demand Distance Vector Routing. ACM Mobile Computing and Communications Review (MC2R), Vol 6. No. 3, pp. 106-107, July 2002.
[35] Manel Guerrero Zapata, Asokan N. Securing Ad-hoc Routing Protocols, In Proceedings of the 2002 ACM Workshop on Wireless Security (WiSe 2002), pages 1-10. September 2002.
[36] Papadimitratos P, Haas Z J. Secure Link State Routing for Mobile Ad hoc Networks, IEEE Workshop on Security and Assurance in Ad hoc Networks, In Conjunction with The 2003 International Symposium on Applications and the Internet, Orlando, FL, January 28, 2003.
[37] Haas Z J, Pearlman M R. The Performance of query control schemes for the Zone Routing Protocol, ACM/IEEE Trans. net, 2001, 9(4): 407-438.
[38] Hu Y C, Perrig A, Johnso D B. Packet Leashes: A Defense against Wormhole Attacks in Wireless Networks, in Proceedings of the Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies(INFOCOM 2003), IEEE, San Francisco, CA, April, 2003.
[39] Zhang Y G, Lee W K. Intrusion Detection in Wireless Ad-hoc Networks, in Proceedings of The Sixth International Conference on Mobile Computing and Networking (MobiCom 2000), Boston, MA, August 2000.
[40] Zhang Y G, Lee W K. Intrusion Detection Techniques for Mobile Wireless Networks, Mobile Networks and Applications, 2003.
[41] Oleg Kachirski, Ratan Guha, Intrusion Detection Using Mobile Agents in Wireless Ad hoc Networks, IEEE Workshop on Knowledge Media Networking (KMN'02).
[42] Tseng C Y, Poornima Balasubramanyam, Calvin Ko, et al. A Specification-Based Intrusion Detection

System For AODV,2003 ACM Workshop on Security of Ad hoc and Sensor Networks (SASN'03) October 31 2003,George W. Johnson Center at George Mason University,Fairfax,VA,USA.
[43] Michiardi P,Molva R. Simulation-based Analysis of Security Exposures in Mobile Ad hoc Networks, in Proceedings of European Wireless Conference 2002.
[44] Levente Buttyan,Jean-Pierre Hubaux. Enforcing Service Availability in Mobile Ad-hoc WANs,In Proceedings of the IEEE/ACM Workshop on Mobile Ad hoc Networking and Computing (MobiHOC),Boston,MA,USA,August 2000.
[45] Zhong Sheng,Chen Jiang,Yang Richard Yang. Sprite: A Simple,Cheat-proof,Credit-Based System for Mobile Ad-hoc Networks,in Proceedings of the Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003), IEEE, San Francisco, CA, April 2003.
[46] Sergio Marti,Giuli T J,Kevin Lai,et al. Mitigating Routing Misbehavior in Mobile Ad hoc Networks. In Proceedings of the Sixth International Conference on Mobile Computing and Networking (Mobicom),Boston,August 2000.
[47] Sonja Buchegger, Jean-Yves Le Boudec. Performance Analysis of the CONFIDANT Protocol: Cooperation Of Nodes-Fairness In Distributed Ad-hoc NeTworks In Proceedings of IEEE/ACM Workshop on Mobile Ad hoc Networking and Computing (MobiHOC), EPFL Lausanne, Switzerland,June 2002.
[48] Sonja Buchegger,Le Boudec J Y. Nodes Bearing Grudges: Towards Routing Security,Fairness,and Robustness,in Mobile Ad hoc Networks 10th Euromicro Workshop on Parallel, Distributed and Network-based Processing,Canary Islands,Spain,January 2002. IEEE Computer Society.
[49] Pietro Michiardi. Refik Molva. Core: A COllaborative REputation mechanism to enforce node cooperation, in Mobile Ad hoc Networks in Communication and Multimedia Security 2002 Conference.
[50] Kong J J,Haiyun Luo,Kaixin Xu,et al. Adaptive Security for Multilevel Ad hoc Networks,Wireless Communications and Mobile Computing,2002,2(5): 533-547.
[51] Kaya T,Lin G,Noubir G,et al. Secure Multicast Groups on Ad hoc Networks In Proceedings of the 2003 ACM Workshop on Security of Ad hoc and Sensor Networks (SASN'03),October 31,2003 George W. Johnson Center at George Mason University,Fairfax,2003.
[52] Loukas Lazos, Radha Poovendran. Energy-Aware Secure Multicast Communication in Ad-hoc Networks Using Geographic Location Information, IEEE International Conference on Acoustics Speech and Signal Processing,2003,6-10.

CHAPTER 3

第3章　无线自组织网络安全架构

摘要：无线自组织网络是一种完全由移动主机构成的网络，它具有自组织无中心的特点。这些特点使得常规安全系统难以应用。本文借鉴免疫系统的思想，采用多代理来模拟实现淋巴细胞的免疫功能，设计了一个适用于无线自组织网络的安全架构，实现了对整个网络的入侵检测和入侵响应，同时还具有学习机制、分布式、自适应等特点。最后，通过实例分析阐明了安全架构的有效性。

关键字：无线自组织网络、网络安全、人工免疫系统、移动代理、入侵检测。

3.1　引言

无线自组织网络作为一种新型的移动多跳无线网络[1]，与传统的无线网络不同，它不依赖于任何固定的基础设施和管理中心，而是通过传输范围有限的移动节点间的相互协作和自我组织来保持网络连接和实现数据的传递。

无线自组织网络的特点对其安全方案提出一些独特的要求：

(1) 不但要抵抗外界攻击，而且能阻止内部的进攻。无线自组织网络的一个主要特点是节点的移动性，网络如果布置于战场环境或其他未受保护的地域，则可能造成节点落入敌方或攻击者的手中，节点的一些资料，如密钥等，就会被攻击者获得。这样就会导致攻击者以网络成员的身份从内部发起进攻。

(2) 安全方案必须是分布式的。无线自组织网络本身就是自组织的，没有中心控制节点，增加任何集中的控制或安全服务器都将导致单点失败。

(3) 发现入侵者后，应及时将其孤立起来并排除出网络，以减少对网络的攻击。以上这些要求使得一些常规的安全策略不能很好地发挥作用。

现阶段在无线自组织网络安全技术研究主要分为四个方面：

(1) 密钥的管理与认证。研究在无线自组织网络中无中心自组织的情况下，如何实现密钥的分配，如何实现相互认证，主要采用两种方式，基于门限密钥的管理方案[2]和基于PGP的自组织的认证[3]。

(2) 路由安全方案，研究如何对路由协议提供安全技术，对路由协议中的报文提供完整性、认证等安全保障，防止恶意篡改的发生[4,5]。

(3) 入侵检测,研究如何在网络运行中及时发现恶意节点的入侵,通常采用邻居监视,合作检测的方法[6]。

(4) 增强合作,研究防止节点为了节省自己的能源,不参与网络交换,主要采用两种机制:基于惩罚的机制,利用邻居监视,不参与网络交换将被排除出网络。基于奖励的机制[7],节点转发报文就给与虚拟货币作为奖励,只有持有货币才能发送自己的报文[8]。上述针对无线自组织网络的安全技术存在一些不足之处,前两种技术,密钥的管理和路由安全只能被动阻止入侵,不能发现入侵者,第三种技术入侵检测,虽然能发现入侵者,但不能清除入侵者。上述方案都没有任何学习机制。

人体的免疫系统是人体的安全保障系统,是一个高度进化的生物系统,它旨在区分外部有害病原体(抗原)和自身组织,从而清除抗原并保持有机体的稳定。从信息处理的角度来看,生物免疫系统是一个高度并行、分布、自适应和自组织的系统,具有很强的学习、识别、记忆和特征提取能力,不但能够识别抗原而且清除抗原,维护人体整个系统功能的正常运转。网络安全系统也应是一个能够识别外界入侵者,保护整个网络正常运行的系统。两者具有很多相似性,因此,本文借鉴人体免疫系统的思想,针对无线自组织网络设计一个分布、自适应的安全系统架构。该安全架构不仅实现入侵检测,而且能够产生入侵响应,最终将入侵者排除出网络之外。本章的主要创新之处为将免疫机理引入到安全架构设计之中,提出基于免疫机制的安全架构。

Forrest[9]、Hofmeyr 和 Forrest[10,11]将免疫机制引入计算机领域,提出了一个分布式的、健壮的、自适应的人工免疫系统架构,并将其应用于网络入侵检测。在该系统中,免疫细胞被定义为检测集。检测集是由一定数量的定长字符串组成。检测是通过 r 个连续们字符串匹配来实现的。通过文献[12]的阴性选择算法来挑选检测集。用于入侵检测时,检测集是源地址、目标地址和端口号所形成的三元组。正常的网络连接三元组为自我,非正常的网络连接为非我,通过学习训练后可检测非正常连接来识别入侵。该系统主要适用于局域网内的入侵检测。Kim 和 Bentley[13,14]提出了一个基于免疫机制的入侵检测模型,该模型由三种算法组成:阴性选择、克隆选择和基因库进化。Dasgupta[15]提出一个基于代理的免疫入侵检测系统,在其系统中,具有各种免疫功能的代理在网络中到处移动,合作监视网络运行,并检测已知和未知的入侵。J. S. Balasubramaniyan 等人提出了一个基于多代理的安全架构——AAFID[16]。它是一个层次性结构,底层是代理负责收集各种网络信息,然后发给上一层的收发器,收发器负责管理代理并处理从代理发来的信息。最上层是监控器,它收集从各个收发器发来的信息,综合形成整个网络的信息,并与用户交互。该方案的优点是采用多代理层次结构,能够灵活配置。它的缺点是对信息的处理仍然是集中式的,会导致大量的通信负载和处理延迟。上述几种方案主要是在固定网络的环境下实现的,不适用于无线自组织网络的运行环境。

本章借鉴免疫系统的思想,通过不同的代理来实现各种免疫细胞的机制,如监视、匹配、响应等,各种代理之间既独立工作又相互联系,它们无需外界的控制,自组织地实现了对整个网络的安全监控与响应,同时具备免疫系统的一些特点,如学习、识别和记忆性,并行和分布性,可扩展性等。本章其余部分如下安排,3.2 节简述免疫系统和代理的原理及特点,3.3 节论述了安全架构的组成及各种代理的结构和功能,3.4 节进行了实例分析。3.5 节讨论了安全架构的特点,3.6 节为结束语。

3.2　免疫系统及移动代理概述

3.2.1　免疫机理

本节简要介绍免疫的基本原理，主要内容来自于文献[17～20]。免疫系统是一种人或动物身体防御外界病原体入侵的复杂系统，其主要功能是识别体内组织的细胞，将其归类为“自我”和“非我”，并引发适当的防卫机制去除“非我”。自我对应于机体自身的组织；非我对应于外来有害病原或者体内病变组织，也称抗原。免疫应答是机体免疫系统与抗原相互作用的表现形式，指免疫活性细胞对抗原的识别、活化、增殖与分化以及产生抗体，并与抗原作用，将其破坏、清除的过程。

免疫系统由固有免疫和获得性免疫组成。固有免疫是人体抵抗外界病原体入侵的第一道防线，它由皮肤和吞噬细胞组成。固有免疫为机体先天获得，可抵抗一些常见病原体，同时为获得性免疫应答赢得时间。获得性免疫是第二道防线，它主要由淋巴细胞组成。它能够识别特定的抗原，并产生针对性的免疫反应，并能够记忆以前遇到过的抗原，在以后如果有类似抗原入侵体内，会产生快速而有效的反应。淋巴细胞又分为B细胞和T细胞两种。B细胞的主要功能是产生抗体，与抗原结合并将抗原清除出体外。T细胞的主要功能是识别抗原并调节其他细胞对抗原实施攻击。

免疫识别是免疫系统的主要功能，识别的本质是区分“自我”和“非我”。免疫识别是通过淋巴细胞上的抗原识别受体与抗原的结合实现的。免疫识别过程同时也是一个学习的过程，学习的结果是免疫细胞能够将抗原的特征提取，并且最优个体以免疫记忆的形式得到保存。免疫学习大致可分为两种：一种发生在遇到新的抗原阶段，即免疫系统首次识别一种新的抗原时，其应答时间相对较长，淋巴细胞需要一定的时间进行调整以更好地识别抗原，并在识别结束后以最优抗体的形式保留对该抗原的记忆信息。而当免疫系统再次遇到相同或者结构相似的抗原时，在联想记忆的作用下，其应答速度大大提高。并且产生针对性的抗体去除病原，这个过程是一个增强式学习过程，对应于再次应答。

免疫响应可分为三个阶段：

(1) 免疫识别阶段。当外界抗原进入体内时，首先进行由淋巴中的T细胞进行免疫识别，识别是通过淋巴细胞上不同的抗原识别受体与抗原的结合来实现的。如果以前识别过类似抗原并保留了信息，T细胞的记忆受体就会很快与抗原结合，实现免疫识别。

(2) 抗体产生阶段。如果T细胞识别了抗原，它会被激活并产生激素，去激活B细胞繁殖并产生大量抗体。

(3) 抗原消除阶段。B细胞产生的抗体与抗原结合，破坏了抗原的活性，使其分解排出体外。抗原的刺激信息消失，免疫响应结束，免疫系统恢复到正常状态。

3.2.2　移动代理简介

移动代理是为了达到某个特定的目标，在对外部环境的相互作用基础上，通过对环境状态的认识以及和其他代理的协作，自主地推进问题解决的处理单位。移动代理具有以下特点：

(1) 自主性。代理拥有内部自治机制和问题解决机制,能够控制自己的行为和内部状态。无需外界的指令即可根据自己的知识和收集到的信息进行判断和行为。

(2) 协作性。代理之间相互通信合作,共同完成某项任务。

(3) 反应性。代理能够根据网络环境的变化做出适当的调整,具有自适应能力。

(4) 移动性。代理能够自主地通过网络从一台主机移动到另一台主机。

从移动代理的特点来看,它具有免疫系统中淋巴细胞的基本特征,用移动代理来模拟免疫细胞,更容易实现其免疫功能。

3.3 安全架构

3.3.1 总体结构

在免疫系统中,淋巴 T 细胞、B 细胞以及胸腺、骨髓形成一个复杂的系统。胸腺产生的 T 细胞和骨髓产生的 B 细胞随血液在全身循环流动,监视机体组织情况,一旦有病原体入侵时,T 细胞发现抗原并激发 B 细胞产生抗体,最终抗体与病原体结合并清除于体外。我们可利用不同代理来模拟实现不同的淋巴细胞。相对于不同类型的淋巴细胞,我们将系统中的代理分类为监视代理、决策代理和攻击代理。

监视代理驻留在每一个节点内监视其邻居节点的行为并将其监视所获得的信息编码后发往决策代理。决策代理收集区域内每个监视代理收到的发来的监视信息,将其综合,形成某个节点一段时期内的行为序列,再与免疫记忆和路由协议进行对比,做出判断。如果发现入侵者,决策代理将产生大量攻击代理。这些攻击代理移动并驻留入侵者周围的节点,形成一道防火墙将入侵者包围隔离,同时将其链路断开,阻止其任何报文的发送与接收。图 3-1 显示总体结构的三个关键组成部分:监视代理、决策代理和攻击代理,以及它们相互之间的关系。

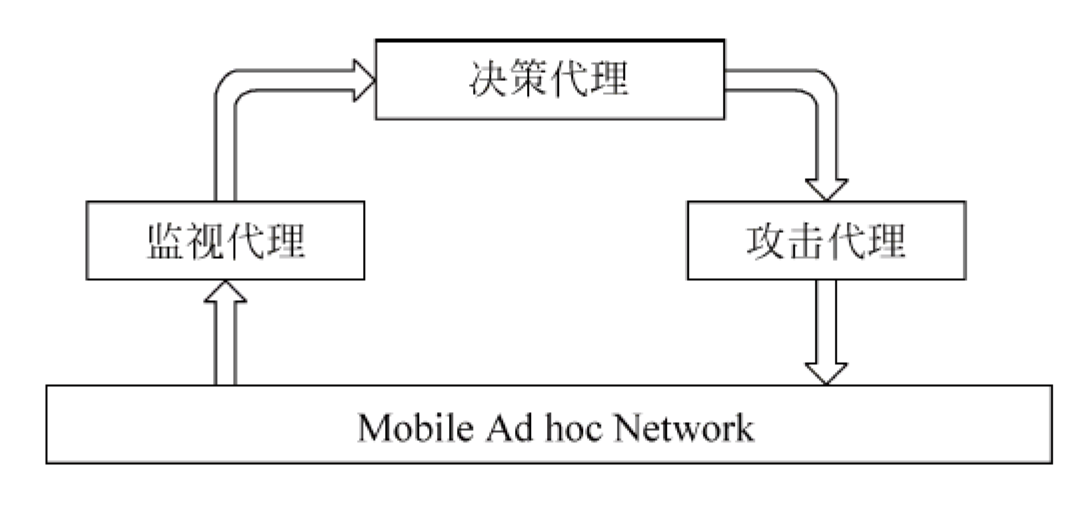


图 3-1 系统结构图

3.3.2 监视代理构成

监视代理类似于免疫系统中 T 细胞,它位于每个节点上,负责收集其周围邻居节点的行为信息并发送给决策代理。对于新加入节点,邻居节点的监视代理复制一份并将其移入新节点。监视代理内部有监听、过滤、编码和通信四个功能模块,图 3-2 显示其基本结构。

监听模块负责收集其所能收到所有邻居节点的通信内容。因为无线通信是无方向性的,任何在其通信范围内的节点均可收到其信息,所以双方的通信能够被第三方监听。但是

监听所收到的信息量很大,不可能也没必要将所有信息都发往决策代理。首先过滤模块对监听收到的初始信息进行过滤,将一些不必要的信息进行滤除,比如用于节点之间保持连接的 hello 报文就可过滤掉。其次,编码模块负责对过滤后的重要信息进行压缩编码。这样可大大减少代理之间的通信量。可采用下列数字对网络行为进行编码。

- 1 代表邻居节点发送 RREQ;
- 2 代表邻居节点收到 RREQ;
- 3 代表邻居节点发送 RREP;
- 4 代表邻居节点收到 RREP;
- 5 代表邻居节点发送 RRER;
- 6 代表邻居节点收到 RRER;
- 7 代表邻居节点发送 DATA;
- 8 代表邻居节点收到 DATA。

图 3-2 监视代理的结构

节点对周围邻居节点行为编码后记录下来,例如,某个节点对邻居节点 A、B、C 的行为记录如下:

A:21214343 B:87878787 C:2244688

通信模块将这些编码后监视信息发送到决策代理。

3.3.3 决策代理构成

决策代理类似于免疫系统中 B 细胞,它是整个安全架构的核心,担负着信息的收集、判断、攻击代理产生等任务。与此类似,免疫系统的淋巴细胞分布在全身并随血液在全身循环流动,对整个机体进行监视。决策代理驻留在节点上,分布于 Ad hoc 网络的各处并随 Ad hoc 节点移动对整个网络进行监视。为了减少对系统资源的占用,决策代理不需要驻留在每一个节点上,它只是平均分布于网络中,即对整个无线自组织网络按区域划分,每个决策代理负责一个区域的监控。决策代理的结构可分为三个模块两个数据库,其相互关系如图 3-3 所示。

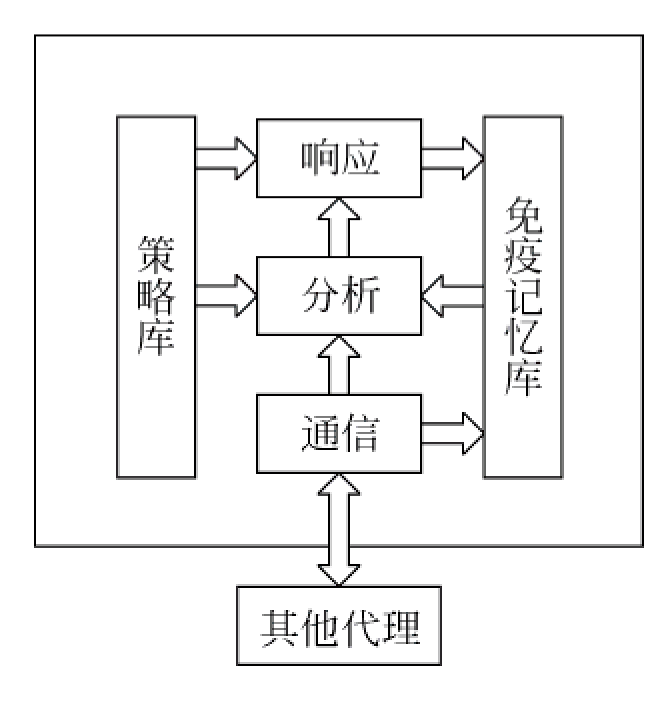


图 3-3 决策代理的结构图

通信模块,用于与监视代理进行通信,收集监视代理对邻居节点的监视信息,还可用与其他决策代理进行通信,交换免疫记忆信息。分析模块,用于对监视代理发来的各节点信息进行综合判断,首先与免疫记忆库的中信息进行对比,如果符合以前曾遇到过入侵者行为模式,就可直接断定为入侵者,这里采用的是基于特征的入侵检测方法,检测效率和准确率都比较高。如果与免疫记忆库中的入侵特征不相匹配,下一步采用基于规范的入侵检测方法,从策略库取出路由协议规范,对节点行为进行判断,该方法能检测出未知的攻击行为,并将其特征存入免疫记忆库。如果某个节点的行为只有少数不正常,则不一定是入侵节点,可能是线路故障,只有超过一定的次数才判定为入侵节点。免疫记忆库中存储一些入

侵节点的行为特征,它有两个来源,一是本节点的以前的一些判断结果,二是通过交换从其他决策代理的免疫记忆库取来的信息,它的作用与免疫系统一样,为第二次遇到类似入侵时,提供快速的识别响应。图 3-4 显示整个处理过程。

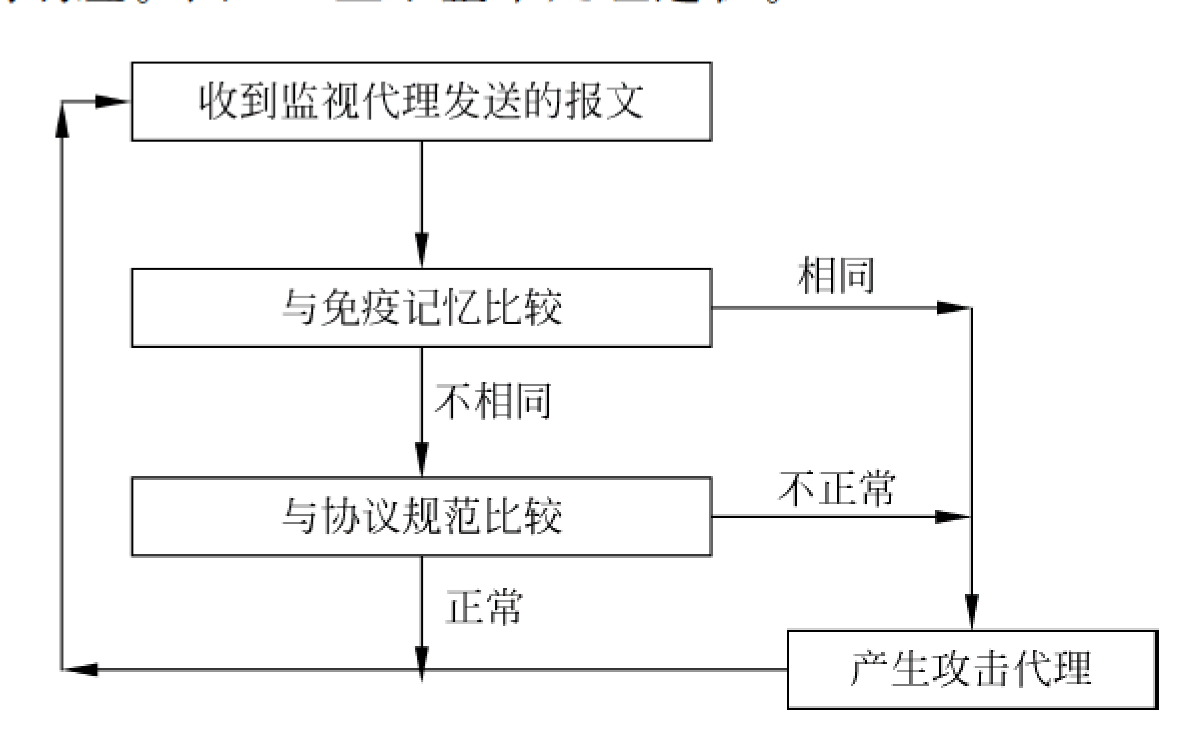


图 3-4 决策代理处理流程

因为无线自组织网络的节点拓扑是动态变化的,无法建立固定的某个监视代理向决策代理汇报的关系,同时为了节省有限的网络无线信道资源,监视代理也不能采用泛洪发送监视报文的方式。我们采用查询应答方式,每个监视代理只是被动地监视邻居节点的行为并将其记录下来,它不主动发送报文,只是等待决策代理的查询报文。决策代理定时向周围节点广播查询报文,该报文的传输范围由决策代理的监视范围所定,例如:以决策代理为中心一跳或几跳的通信范围。当监视代理收到查询报文后将监视收集的信息发回决策代理。决策代理收到各监视代理的报文后进行分析。因为对区域内每个监视代理的信息的综合,所以对某个节点的监视就会比较完全,不会有信息遗漏。

对节点的行为分析,主要基于路由协议的规范,采用基于规范的入侵检测方法,如正常节点在接收到发往其他节点的报文应该及时转发出去。在上节的监视信息中 B 节点为 878787,说明其接收报文后即进行了转发,是正常网络操作。C 节点为 2244688,表示 C 节点只收报文不转发报文,不是正常网络行为。当某个节点的不正常行为超过了一定的限度,就可认定为其为恶意节点。然后将其行为的特征存入免疫记忆库,以后再遇到同类行为时,就可及时发现。

由于网络节点的动态特性,某个区域的决策代理可能由于节点移动、节点退出而空缺或决策代理遭到攻击而失效。这时,监视代理就收不到决策代理的查询报文,当超过了一定的时间限度,就可推断该区域的决策代理已经不存在。这时该区域的节点选举一个节点驻留决策代理,并由该节点从周围节点请求一个决策代理。该请求报文达到周围区域的某个决策代理时,该决策代理复制一份,复制后的决策代理移动到请求节点。可能会有多个决策代理响应,最先到达的决策代理发挥作用,随后到达的抛弃。选举算法可采用竞争方式,谁先申请谁得到;也可采用协商方式,哪个节点资源充足,哪个就作为决策代理的驻留节点。

3.3.4 攻击代理构成

当免疫系统发现有外界病原体进入体内时,淋巴 B 细胞会产生大量抗体与抗原结合,

清除抗原，将其排除体现外。攻击代理就担任淋巴系统中的抗体的角色，主要负责将恶意节点包围并孤立。

无线自组织网络中节点必须通过邻居节点的转发才能参与网络，攻击节点代理移动到恶意节点的周围邻居节点上，将恶意节点包围起来，不再接受恶意节点的路由请求和报文发送，也不再向其转发报文。虽然恶意节点在网络中，但它不能参与任何网络功能，等于将它排除于网络，这样才能够最大限度减少对网络的危害。攻击代理有定位、移动、隔离和自杀四个功能模块，其结构关系如图 3-5 所示。

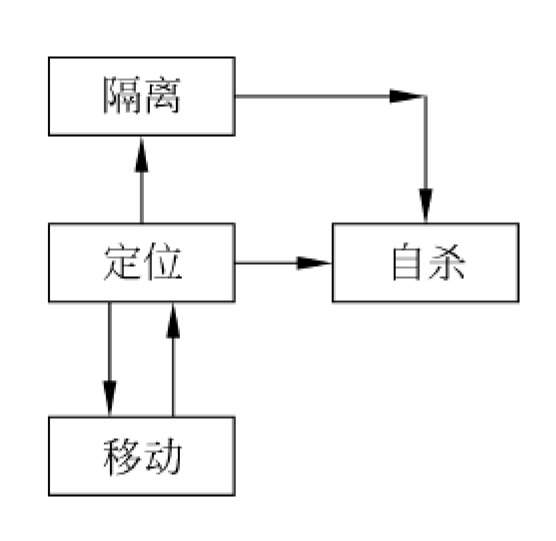


图 3-5　攻击代理结构图

当决策代理产生大量攻击代理时，已经能确定有入侵者就在附近，但不一定就处于邻居的位置，也许相邻一个或两个节点。这些攻击代理需要定位入侵者的位置，并移动到入侵者周围将其包围隔离。定位模块用于搜寻入侵者的位置，指示攻击代理移动的方向。移动模块按照定位模块指出的方向，负责攻击代理移动到入侵者的附近。当攻击代理到达入侵者的邻居节点时，隔离模块将切断入侵者与外界的联系。自杀模块负责结束攻击代理本身的生命，它在两种情况下自杀模块发生作用，一是在一定时间内无法定位到入侵者时，二是入侵者死亡不再需要攻击代理隔离时。自杀模块的作用是防止攻击代理大量长时间占用节点资源，只有当入侵者存在时，才需要攻击代理隔离，入侵者死亡，攻击代理也应自行消失。

3.4　实例分析

为了进一步阐述设计的安全架构的功能和特点，在本节中举一个例子来说明整个入侵检测和响应的过程。图 3-6 为一个无线自组织网络的拓扑图。

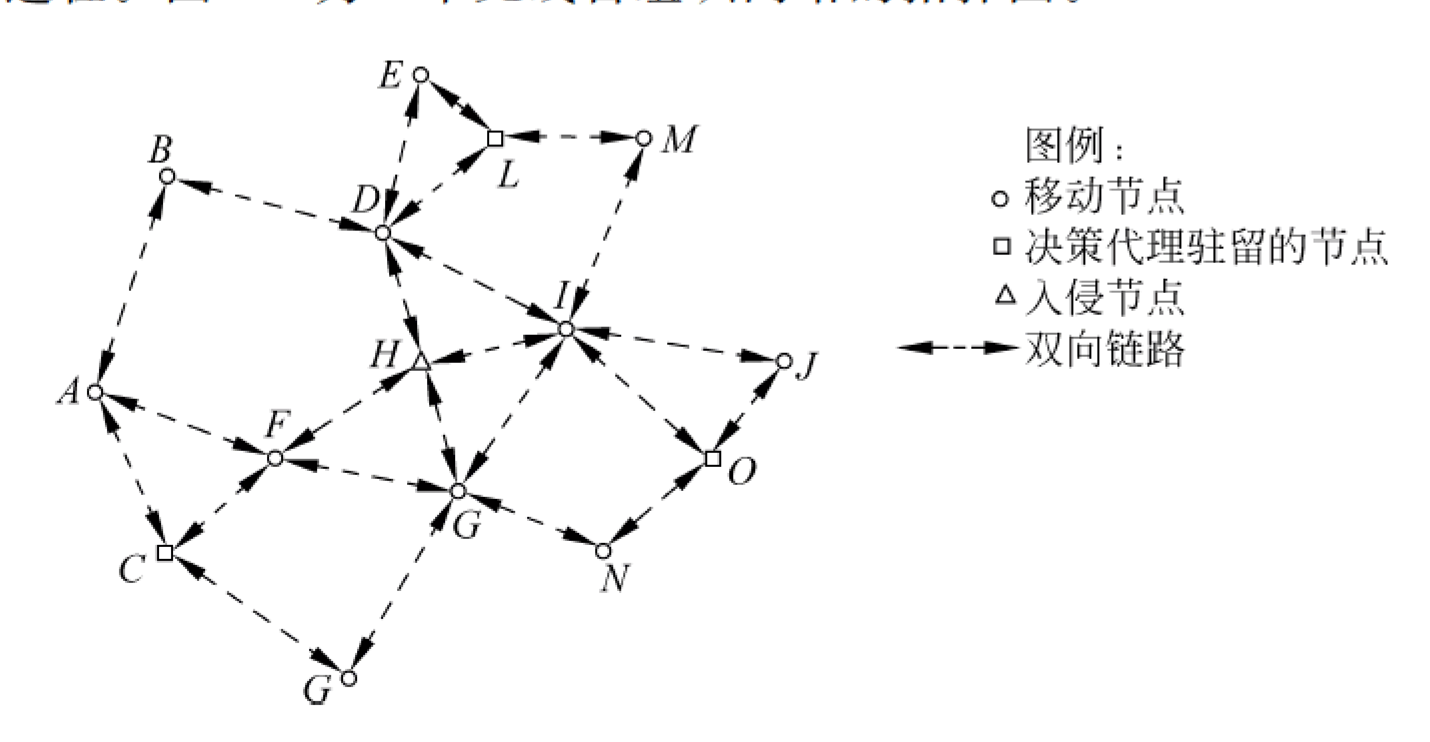


图 3-6　一个无线自组织网络拓扑图

图 3-6 中有 15 个移动节点，从节点 A 到节点 O，相邻节点通过双向链路进行连接，其中节点 H 为入侵者，每个节点都驻留监视代理，监听并收集其邻居节点的行为信息，三个节点 C、L、O 中驻留了决策代理，负责其区域内信息的汇总与决策，如节点 L 上的决策代理就负责汇总节点 D、E、M、L 节点上监视代理所收集的信息。

图 3-7 显示入侵节点 H 开始发动拒绝服务攻击，它向整个网络泛洪发送大量无用数据

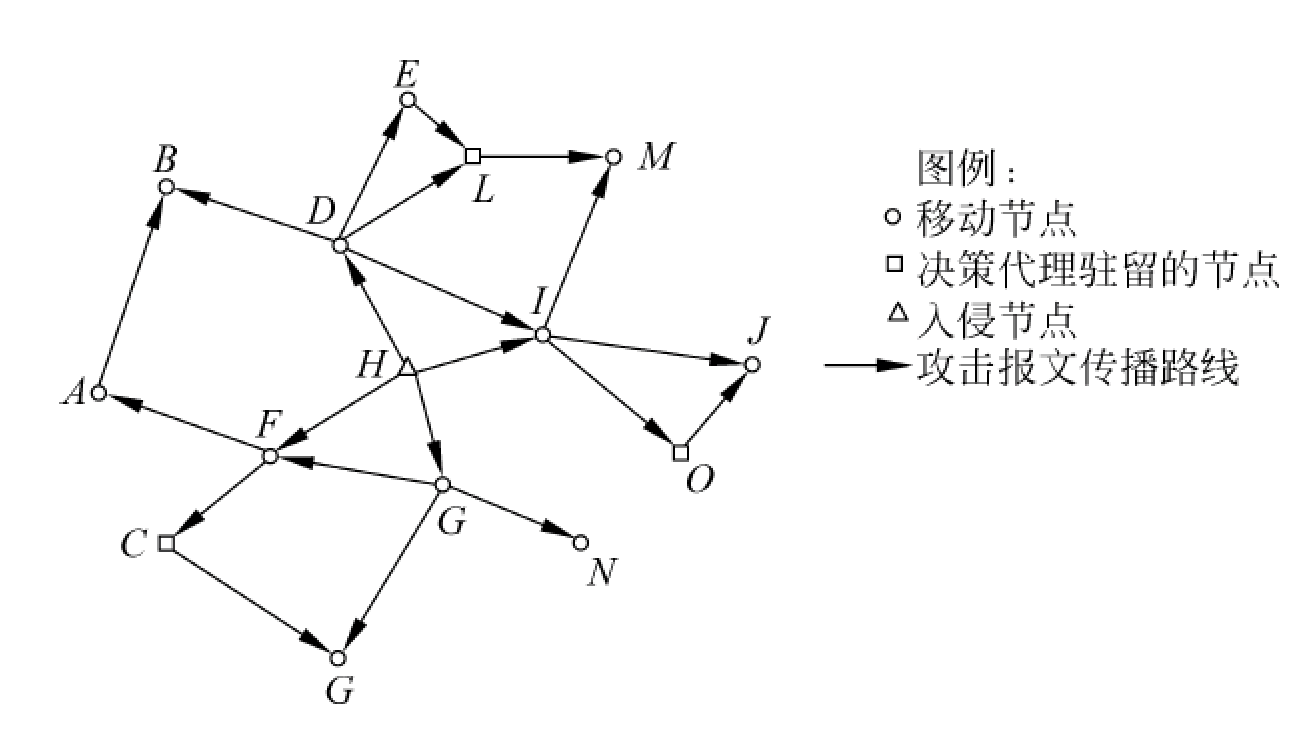


图 3-7 入侵节点发动攻击

报文或路由查询报文，数据报从入侵者周围节点开始向整个网络扩散，大量占用和消耗网络资源，导致其他节点无法正常传送报文。

整个安全架构的响应可分为入侵检测和入侵响应两个过程，首先是入侵检测，节点 F、G、I、D 是 H 的邻居，在节点 F、G、I、D 上的监视代理时刻监视节点 H 的行为并将其行为进行编码，当 H 节点连续发送查询报文时，其行为的编码为 6666666，F 节点上的代理将编码发往 C 节点上的决策代理，D 节点上代理的监视数据发向 L 节点，G、I 节点上代理的监视数据发往 O 节点。决策代理首先在免疫记忆库中进行特征匹配，如果以前有过相同的入侵行为并将其特征存入了免疫记忆库，则可快速匹配和识别入侵。如果匹配不成功，则调用策略库中的路由规范进行判断。判断为入侵行为后，下一步进行入侵响应，决策代理的响应模块开始产生攻击代理。在节点 C、L、O 上的决策代理判断有入侵后，分别产生攻击代理。节点 C 上决策代理产生的攻击代理沿 CF 链路到达入侵者 H 的邻节点 F，到达后将节点 F 与入侵者 H 的链路 FH 中断，拒绝 H 节点的任何路由报文。同样，节点 L 和 O 上的决策代理产生的攻击代理分别到达入侵者的另外三个邻居节点 D、I、G，同时将其与节点 H 的链路 DH、IH、GH 断开。这样入侵者 H 虽然在网络中，但已完全被其周围节点隔离。如图 3-8 所示，攻击代理移动到入侵者周围四个节点驻留，形成一道移动防火墙，如图中的虚线，将入侵者隔离。图 3-8 显示攻击代理的移动和孤立的过程。所谓移动防火墙，是指由节点 F、G、I、D 所构成的阻止入侵者 H 的防火墙是动态的，它会随着入侵者 H 的移动而移动。

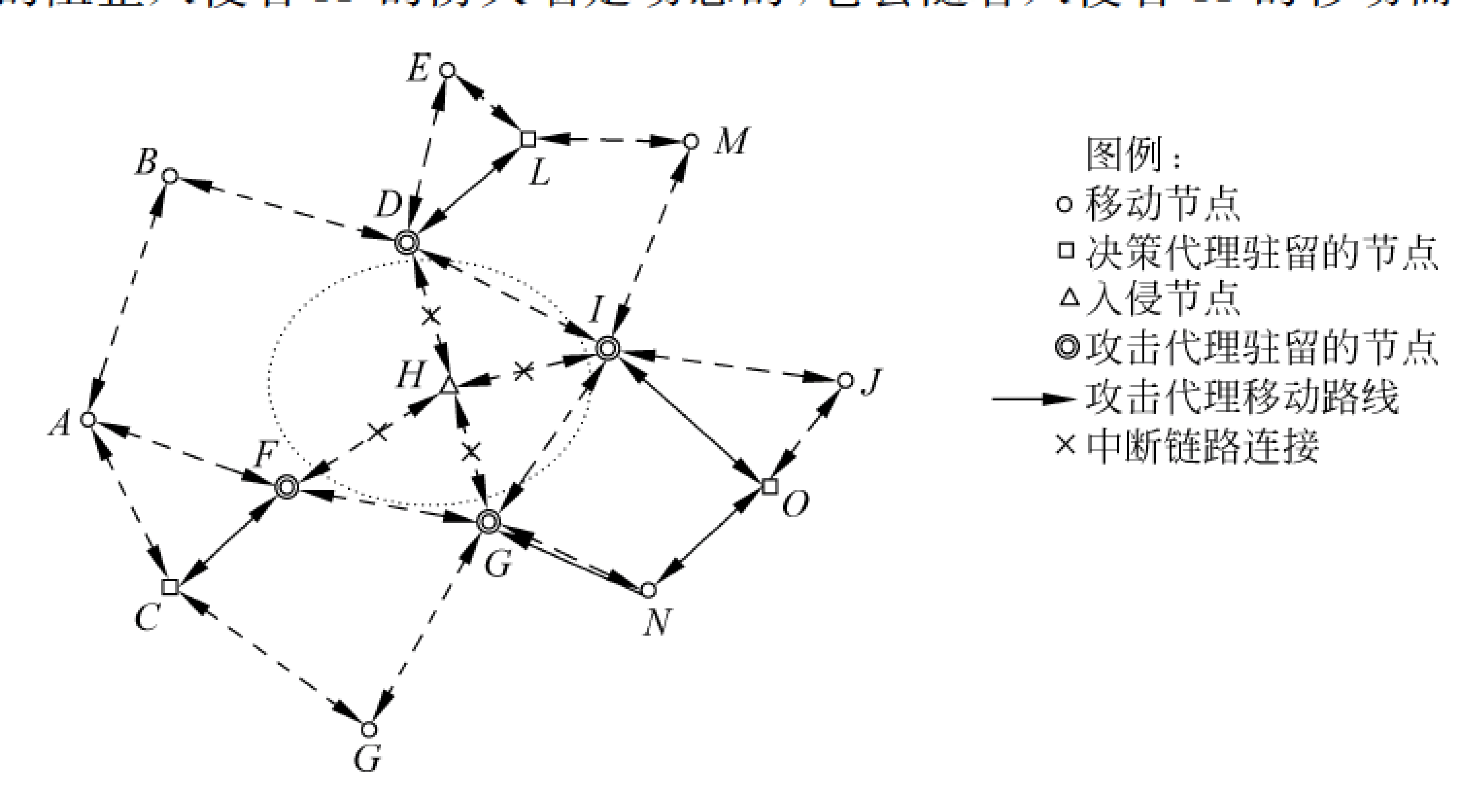


图 3-8 攻击代理包围并孤立入侵节点

从上述分析可以看出，遍布整个网络的监视代理实现对每个节点的监控，将节点的行为编码后发送到决策代理，决策代理根据免疫记忆和安全规范进行判断。如果发现入侵者，则决策代理产生攻击代理，由攻击代理将入侵者包围并隔离，最终消除入侵的影响，实现网络的正常运转。

3.5　安全架构的特点

可将无线自组织网络比作整个机体，利用不同的移动代理来模拟实现免疫系统中各种淋巴细胞的功能和作用，实现一个具有免疫特征的安全架构，安全架构组件与免疫系统组织对应关系如表 3-1 所示。

表 3-1　免疫系统与安全架构的映射表

免 疫 系 统	安 全 架 构	免 疫 系 统	安 全 架 构
整个机体	整个无线自组织网络	B 细胞	决策代理
机体组织细胞	无线自组织网络中正常的节点	抗体	攻击代理
抗原	无线自组织网络中入侵节点	结合并清除抗原	攻击代理包围并孤立入侵节点
淋巴细胞	移动代理	免疫记忆	免疫记忆库
T 细胞	监视代理		

借鉴免疫机制设计的安全架构，因而它也具有免疫系统的一些重要特点。

（1）分布式和自治性：免疫细胞分布于全身各处，它们每个淋巴细胞根据本地的实际情况，做出不同的反应。既没有也不需要全局的控制中心，这就意味着没有单点失败的危险。在安全架构中，也没有集中的控制，分布于整个网络的代理既相互联系又各自独立地运行着，决策代理根据本地监视代理收集的信息，自主地做出判断，不需要外界的任何干预。这非常适合于 Ad hoc 网络自组织无中心的特点，也是我们安全架构区别于其他安全系统的一个重要方面。

（2）学习和自适应性：当免疫系统第一次遇到某种抗原时，需要一段时间进行识别，但它经过第一次识别后，它就会记忆该种抗原的信息，以后再遇到同类抗原时，反应时间就会大大缩短。在安全架构中，同样具有学习机制，对初次认定入侵者的行为能够提取特征，存入免疫记忆数据库，第二次遇到类似入侵者时，就能及时识别并响应。

（3）以行为标识入侵者：免疫系统以基因来标识并识别抗原。在许多安全系统中，节点的区分通过字母或数字标记，如 IP 地址等，这容易导致入侵者通过改变其标识而逃脱监视。在安全架构中，以入侵者的行为来标记并识别入侵者，只要入侵者不改变其行为就不能逃脱。

（4）可扩展性：当无线自组织网络节点数量增加网络扩大时，只是需要在每个新增节点中驻留一个监视代理，再增加一些决策代理。因为监视、判断和响应都局限于本地，所以安全架构对资源占用和计算量并没有明显增加。

3.6 小结

本章首先简述了生物免疫机制的原理。然后借鉴免疫系统的思想，采用多代理来模拟实现淋巴细胞的免疫功能，设计了一个适用于无线自组织网络的安全架构，并系统地论述了安全架构的设计思想、结构、组件、工作流程。最后，通过实例分析阐明了安全架构的有效性。

参考文献

[1] Corson S, Macker J. Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations, RFC 2501, January 1999.

[2] Zhou Lidong, Zygmunt J Haas. Securing Ad hoc Networks, IEEE Networks Special Issue on Network Security, November/December, 1999.

[3] Srdjan Capkun, Levente Nuttyan, Jean-Pierre Hubaux. Self-organized public-key Management for Mobile Ad hoc Networks, IEEE Transactions on Mobile Computing, 2003, 2(1).

[4] Papadimitratos P, Haas Z. Secure routing for mobile Ad hoc Networks, in Proceedings of the SCS communication Networks and Distributed Systems Modeling and Simulation Conference, San Antonio, TX, January, 2002: 27-31.

[5] Hu Y C, Adrian Perrig, David B. Johnson. Ariadne: A secure On-Demand Routing Protocol for Ad hoc Networks. In Proceedings of the MobiCom 2002, September, Atlanta, Georgia, 2002: 23-28.

[6] Zhang Yongguang, Lee Wenke. Intrusion Detection Techniques for Mobile Wireless Networks. Mobile Networks and Applications, 2003.

[7] Sonja Buchegger, Jean-Yves Le Boudec. Performance Analysis of the CONFIDANT Protocol: Cooperation Of Nodes-Fairness In Distributed Ad-hoc NeTworks In Proceedings of IEEE/ACM Workshop on Mobile Ad hoc Networking and Computing (MobiHOC). IEEE EPFL Lausanne, Switzerland, June 2002.

[8] Zhong Sheng, Chen Jiang, Yang Richard Yang. Sprite: A Simple, Cheat-proof, Credit-Based System for Mobile Ad-hoc Networks, in Proceedings of the Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2003), IEEE, San Francisco, CA, April 2003.

[9] Forrest S, Hofmeyr S, Somayaji A. Computer Immunology, Communications of the ACM, 1997, 40(10): 88-96.

[10] Hofmeyr S, Forrest S. Immunity by Design: An Artificial Immune System, In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO), Morgan-Kaufmann, San Francisco, CA, 1999: 1289-1296.

[11] Hofmeyr S, Forrest S. Architecture for an Artificial Immune System, Evolutionary Computation Journal, 2000, 8(4): 443-473.

[12] Forrest S, Perelson A S, Allen L, et al. Self-Nonself Discrimination in a Computer, In Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy, Los Alamos, CA: IEEE Computer Society Press, 1994.

[13] Kim J, Bentley P. The Human Immune System and Network Intrusion Detection, 7th European Congress on Intelligent Techniques and Soft Computing (EUFIT '99), Aachen, September, 1999: 13-19.

[14] Kim J, Bentley P J. Negative Selection and Niching by an Artificial Immune System for Network Intrusion Detection, Genetic and Evolutionary Computation Conference (GECCO '99), Orlando, Florida, July 1999: 13-17.

[15] Dipankar Dasgupta, Immunity-Based Intrusion Detection Systems: A General Framework, In the proceedings of the 22nd National Information Systems Security Conference (NISSC), October 1999: 18-21.

[16] Balasubramaniyan J S, Garcia-Fernandez J O, Isacoff D, et al. An Architecture for Intrusion Detection Using Autonomous Agents, Purdue University Technical Report. 1998.

[17] Somayaji S Hofmeyr, Forrest S. Principles of a Computer Immune System, 1997 New Security Paradigms Workshop, ACM, 1998: 75-82.

[18] Hofmeyr S. An Interpretative Introduction to the Immune System, University of New Mexico, 1999.

[19] de Castro L N, Timmis J I. Artificial Immune Systems as a Novel Soft Computing Paradigm, Soft Computing journal, 7(7), July, 2003.

[20] 肖人彬，王磊. 人工免疫系统：原理、模型、分析及展望. 计算机学报，2002，12.

CHAPTER 4

第4章　无线自组织网络中DoS攻击模型

摘要：无线自组织网络由于其动态拓扑、无线信道以及各种资源有限的特点，特别容易遭受拒绝服务(DoS)攻击。本章分析了无线自组织网络的安全弱点及其导致的各种 DoS 攻击方式。然后研究了按需路由协议 AODV，发现其中的一些弱点，该弱点可能导致无线自组织网络中一种新的 DoS 攻击模型——Ad hoc Flooding 攻击。该攻击主要针对移动无线自组织网络中的按需路由协议，如 AODV、DSR 等。Ad hoc Flooding 攻击是通过在网络中泛洪发送超量路由查询报文，大量地占用网络通信及节点资源，以至于阻塞节点正常的通信。分析 Ad hoc Flooding 攻击之后，提出了邻居阻止的防御策略，即当入侵者发送大量路由查询报文时，邻居节点降低对其报文的处理优先级，直至不再接收其报文。模拟实验证实，通过这种方法能够有效地阻止网络中的 Ad hoc Flooding 攻击行为。

关键字：无线自组织网络、路由协议、网络安全、拒绝服务、Ad hoc Flooding 攻击。

4.1　引言

DoS(Denial of Service)拒绝服务攻击是批攻击者非法占用或消耗大量共享资源，使得系统没有剩余的资源提供给其他合法用户使用，从而导致系统无法继续提供正常服务的一种攻击方式[1]。随着 Internet 的发展，攻击者已从攻击主机系统为主转变到以攻击网络为主。DoS 攻击使网络信息系统的可用性受到严重的威胁。这类攻击的目标是使被攻击的主机系统瘫痪，提供的服务失效，使受攻击的主机用户无法使用网络连接和网络服务。发起 DoS 攻击的攻击者往往利用信息系统或通信协议本身的安全缺陷，运行编制的特定程序，恶意地消耗有限的系统资源，如 CPU 资源、内存资源、磁盘空间和网络带宽等，使目标系统无法提供正常服务，甚至导致系统服务崩溃。

对于有线网络中的 DoS 攻击，国际上专家学者进行了深入研究，提出许多解决方案[2]。但对于无线自组织网络中的 DoS 攻击的研究才刚刚开始[3,4]。无线自组织网络由于其节点移动、动态拓扑、有限的通信带宽和能源、合作的路由等特点，使得 DoS 攻击比较容易实施并且可能导致更加严重的后果。

本章首先介绍了固定网络中的DoS攻击防护方法的研究进展,分析了无线自组织网络的弱点及其导致的各种DoS攻击方式。然后,提出一种新的拒绝服务攻击模型——Ad hoc Flooding攻击。该攻击可以针对所有按需路由协议,导致路由协议不能正常运行。当发动Ad hoc Flooding攻击时,入侵者在短时间内向网络发送过量路由查询报文,占用和消耗节点和网络通信资源,导致网络性能急剧下降,正常的通信无法进行。针对上述Ad hoc Flooding攻击,我们提出了防御方法,邻居节点监视入侵者的行为,当其行为异常,发送过量的报文时就降低对其报文的处理优先级,如果继续攻击,就会停止接收其报文,从而孤立攻击者。本文的主要创新在于提出了Ad hoc Flooding攻击模型及防御方法。

4.2 背景知识

4.2.1 无线自组织网络的安全弱点

传统网络中,主机之间的连接是固定的,网络采用层次化的体系结构,并具有稳定的拓扑。传统网络提供了多种服务以充分利用网络的现有资源,包括路由器服务、命名服务、目录服务等,并且在此基础上实现了相关的安全策略,如加密、认证、访问控制和权限管理、防火墙等。而在无线自组织网络中没有基站或中心节点,所有节点都是移动的,网络的拓扑结构动态变化[5]。并且节点间通过无线信道相连,没有专门的路由器,节点自身同时需要充当路由器,也没有命名服务、目录服务等网络功能。两者的区别导致了在传统网络中能够较好工作的安全机制不再适用于无线自组织网络,主要表现在以下几个方面[6]。

1. 传输信道方面

无线自组织网络采用无线信号作为传输媒介,其信息在空中传输,无需像有线网络一样,要切割通信电缆并搭接才能偷听,任何人都可接收,所以容易被敌方窃听。无线信道又容易遭受敌方的干扰与注入假报文。

2. 移动节点方面

因为节点是自主移动的,不像固定网络节点可以放在安全的房间内,特别是当无线自组织网络布置于战场时,其节点本身的安全性是十分脆弱的。节点移动时可能落入敌手而投降,节点内的密钥、报文等信息都会被破获,投降后的节点又可能以正常的面目重新加入网络,用来获取秘密和破坏网络的正常功能。因此,无线自组织网络不仅要防范外部的入侵,而且要对付内部投降节点的攻击。

3. 动态的拓扑

无线自组织网络中节点的位置是不固定的,可随时移动,造成网络的拓扑不断变化。一条正确的路由可能由于目的节点移动到通信范围之外而不可达,也可能由于路由途径的中间节点移走而中断。因此,难于区别一条错误的路由是因为节点是移动造成的还是虚假路由信息形成的。由于节点的移动性,在某处被识别的恶意节点移动到新的地点,改变标识后,它可重新加入网络。另外由于动态的拓扑,网络没有边界,防火墙也无法应用。

4. 安全机制方面

在传统的公钥密码体制中,用户采用加密、数字签名、报文鉴别码等技术来实现信息的

机密性、完整性、不可抵赖性等安全服务。然而它需要一个信任的认证中心来提供密钥管理服务。但在无线自组织网络中不允许存在单一的认证中心，否则不仅单个认证中心的崩溃将造成整个网络无法获得认证，而且更为严重的是，被攻破认证中心的私钥可能会泄露给攻击者，攻击者可以使用其私钥来签发错误的证书，假冒网络中任意一个移动节点，或废除所有合法的证书，致使网络完全失去了安全性。若通过备份认证中心的方法虽然提高了抗毁性，但也增加了被攻击的目标，任意一个认证中心被攻破，则整个网络就失去了安全性[7]。

5. 路由协议方面

路由协议的实现也是一个安全的弱点，路由算法都假定网络中所有节点是相互合作的，共同去完成网络信息的传递[8]。如果某些节点为节省本身的资源而停止转发数据，这就会影响整个网络性能。更可怕的是投降节点和参与到网络中的恶意节点专门广播假的路由信息，或故意散布大量的无用数据包，从而导致整个网络的崩溃。

4.2.2 无线自组织网络中的 DoS 攻击方式

DoS 攻击是使系统降低或失去其提供正常服务的能力。在无线自组织网络中，DoS 攻击的主要目的是使网络运行混乱，无法进行提出报文传输。由于无线自组织网络的特点，可以有许多发起 DoS 攻击的方式，从物理层到应用层都可以产生 DoS 攻击，但大部分的攻击还是位于网络层。下面介绍一些主要的 DoS 攻击方式。

1. 干扰

干扰是一种简单而有效的物理层 DoS 攻击方式。攻击者不需要加入网络，只要捕捉互网络节点所使用的发送和接收信息的频率，连续发送干扰信号，就能够有效地阻塞节点之间正常通信。对付这种攻击的方法一般采用扩频和跳频通信。

2. 篡改

路由协议假定网络中节点都是相互合作的，转发报文的节点不会修改与其无关的路由信息，所以不检查路由信息的完整性。这使攻击者能够十分容易更改路由信息中任何字段。例如，AODV 路由中的序号和跳数，DSR 路由包中的路由节点序列等，从而产生错误的路由，如重定向、回路等，导致整个网络性能下降。攻击者能够篡改路由报文的根本原因在于节点无法对路由报文进行完整性检测[9]。对付篡改攻击的方法是对整个路由报文或其中的关键信息加入报文鉴别码，节点收到路由报文之后，先进行完整性检测，通过后才能进行处理，以防止非法篡改。

3. 冒充

因为路由协议并不认证报文的地址，所以攻击者可以声称为某个节点加入网络，甚至能够屏蔽某个合法节点，替他接收报文。其根本原因在于节点不能鉴别报文的来源。对付冒充攻击的方法是网络各节点之间实现认证机制，对于节点的行为首先要进行认证，是合法节点才能接收和发送报文。

4. 伪造

攻击者可以伪造并广播假的路由信息。例如：广播某条存在的路由已中断，或编造一条并不存在路由。它可造成回路、分割网络、孤立节点等。其原因在于无法验证报文的内

容。对付伪造攻击方式比较困难，因为要随时掌握整个网络的连通情况，才能辨别某个节点所发出的信息的真假。

5. 资源消耗攻击

攻击者发送大量无用报文，如路由查询报文或数据报文，消耗网络和节点资源，如带宽、内存、CPU、电池等。针对无线自组织网络还有一种的攻击叫剥夺睡眠攻击[10]，攻击者不停发送报文，使移动节点的电源很快耗尽，从而达到拒绝服务的目的。

6. Wormhole 攻击

攻击者在网络中收到报文，通过专用通路传送到另一个攻击者，然后重新发送出去，在两个攻击者之间的通路就叫 Wormhole[11]。如果只是通过 Wormhole 来转发报文，也不会对网络运行产生什么影响。但如果两个串通的攻击者，通过 Wormhole，越过正常的拓扑结构，直接转发路由查询报文，造成错误的路由拓扑信息。图 4-1 为 Wormhole 攻击示意图，从 S 节点到 D 节点的正常路由应该为 S—A—B—C—D，但攻击者 M_1 和 M_2 通过 ABC 建立虚拟专用通道用来转发路由查询报文，这样形成了 S—M_1M_2—D 的路由。因为后者路由跳数少，源节点选择了 S—M_1M_2—D 作为发送路由。

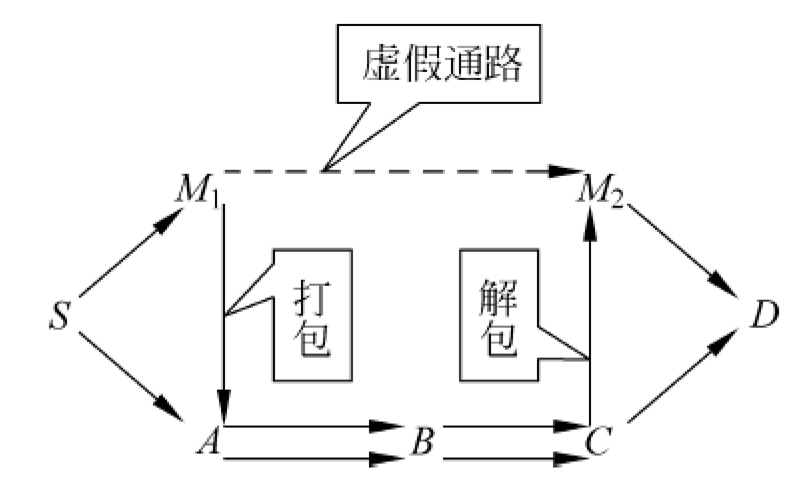


图 4-1　Wormhole 攻击示意图

文献[12]专门提出一种方法 Packet leashes 抵抗 Wormhole 攻击，它基于精确的时间或位置信息来发现并阻止它的攻击。

7. 黑洞攻击

在路由查询或路由更新中，攻击者虚假广播更好的路由，例如更短或更稳定的路由，其目标是阻止建立到目的节点的正确路由。或者在没有至目标节点的路由情况下，抢先宣布有到目标节点的路由，使源节点建立通过该节点的路径，在随后的报文发送中，抛弃通过该节点的报文，形成抛弃报文的黑洞[13]。解决黑洞攻击的一种方法是禁止路由中间节点回答路由查询报文，只能由目标节点回答路由查询。这样虽然在一定程度中阻止了黑洞攻击，但也降低了路由查询的效率。

8. Rushing 攻击

在按需路由协议中，节点路由查询时常采用泛洪查询，可能会导致一个节点收到多个相同的路由查询报文，这时节点只会处理第一个到达的路由查询报文，而将其他相同的路由查询报文抛弃。Rushing 攻击[14]就利用这个弱点，攻击者比其他节点更快地转发路由查询报文，使得其他节点首先收到它转发的报文。它导致两个问题：其一，攻击者短时间内发送大量路由查询遍布整个网络，使得其他节点正常的路由查询无法提交处理而被抛弃；其二，所有建立的路由都通过攻击者。

4.3　相关工作

在固定网络中存在着多种 DoS 攻击方式，如 Smurf、RST flooding、SYN Flooding、ICMP flooding、DNS replay flooding 攻击[15]。其中，SYN Flooding 为主要的攻击方式，它

在所有网络 DoS 攻击中占到 90%～94%。在基于 TCP/IP 协议的网络系统中，为确保通信的另一方能正确无误地接收到信息，建立 TCP 连接时采用了 3 次握手协议。该协议的过程是，首先用户发起建立连接请求，服务器端在收到用户连接请求后，并不是立即建立连接，而是向用户回答该请求并将该请求缓存，等待用户第三次响应到达才建立连接。如果在连接建立的过程中出现了意外情况，导致第三个握手包(TCP ACK)无法到达，缓冲在服务器端的连接请求信息直到超时才会被清除，并释放所占用的资源。

图 4-2 显示一次 TCP 连接建立的过程。SYN Flooding 正是利用这一特性对网络服务器实施攻击。如果大量连接请求数据包发到服务器端后都没有应答，会使服务器端的 TCP 资源迅速枯竭，正常的连接不能进入，甚至可能导致服务器系统崩溃。当攻击者以一个很高的速率给目标发送 SYN 请求，填装某个甚至多个 TCP 连接请求缓存队列。遭到攻击的服务器根据 SYN 请求将 SYN/ACK 发往各处，并等待回应的 TCP ACK 包。由于攻击者的假 SYN 连接请求发送速率足够快，且不发回第三个回复响应，导致服务器的缓存队列迅速被填满。服务器在这种情况下再接收到 SYN 请求连接包时将拒绝该请求，使得真正合法的用户无法连接到该服务器。图 4-3 显示 SYN Flooding 攻击过程。

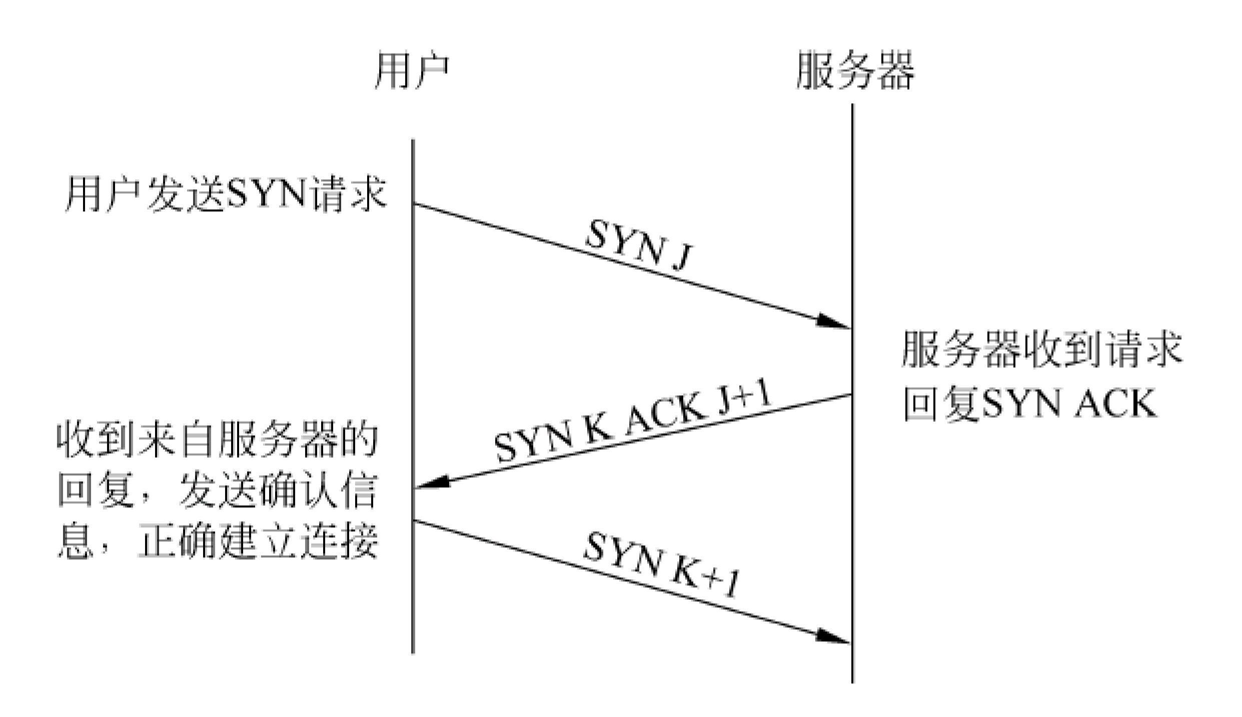


图 4-2　一个 TCP 连接正常建立过程

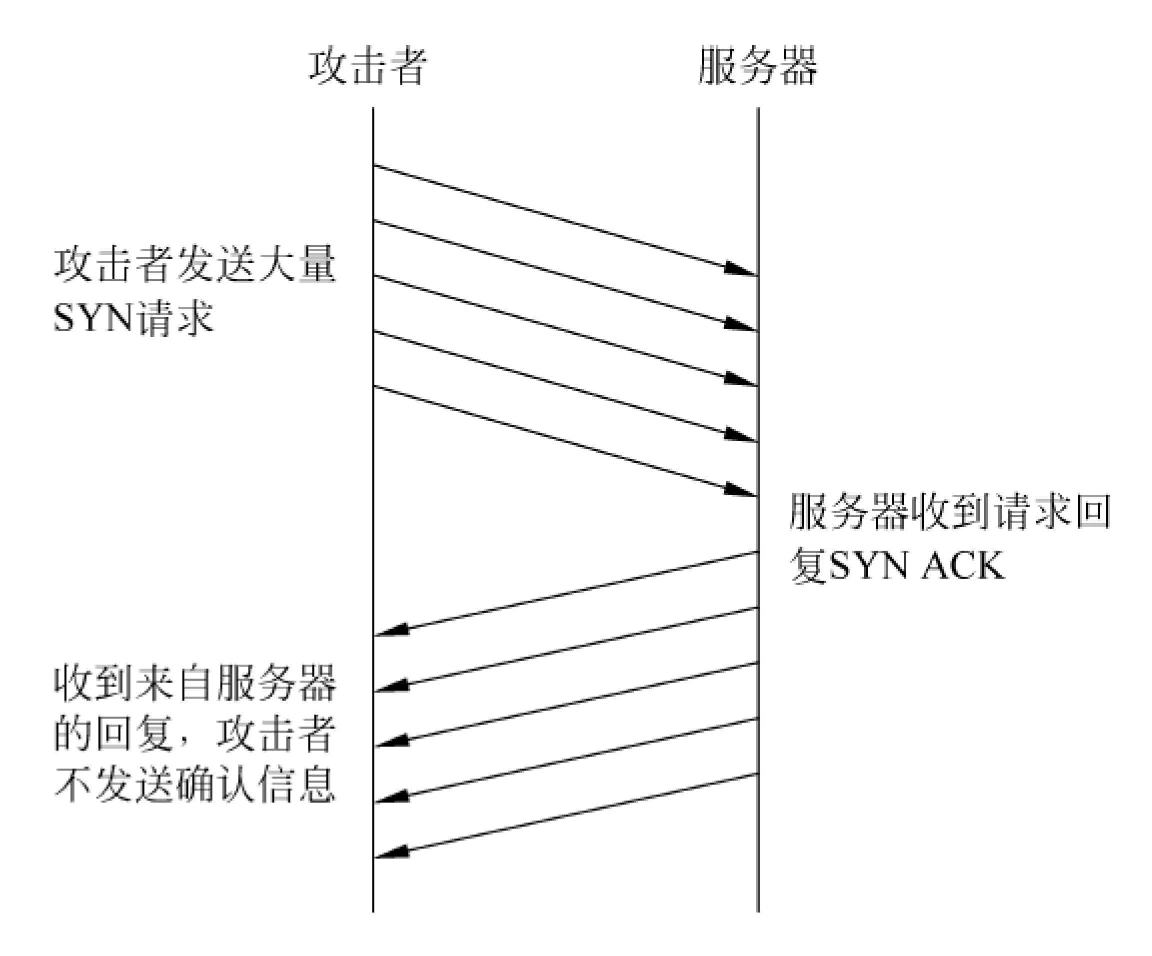


图 4-3　SYN Flooding 攻击过程

由于目前使用的网络协议无法判断正常的 TCP 连接请求和黑客发出的假连接请求。因此有效防范措施应当在系统遭受攻击时仍能够允许正常的 TCP 连接。常见的对 SYN Flooding 的防范措施有基于主机和基于防火墙两类[16~18]，其主要的措施有增大系统连接等待队列的大小、缩短 TCP 连接超时时间、SYN cookies、SYN cache 等。增加系统连接缓冲队列的大小是一种最为简单的防御方法。在受到 SYN Flooding 攻击时，可以起到防御作用。但当攻击者利用分布在网络不同位置上的数百台主机同时发起攻击，连接队列还是会很快被填满，造成系统瘫痪。而且加大等待队列长度将消耗大量系统资源，使系统性能下降。缩短 TCP 连接超时时间是一种防御 SYN Flooding 攻击的有效办法，但也有不足之处，因为对合法用户和非法用户提出的连接请求的超时时间相同。当合法用户提出的连接请求在超时时间内没得到及时服务时，也得不到响应。

基于防火墙和路由器的防御系统在网关处截获 TCP 连接，使用代理的方式与源目标进行连接，只有成功返回时才会发向最终目标。使用基于防火墙的防御措施具有两个优点：服务端不需要改动。保护整个内部网络的主机。但当使用分布式手段对该目标发起 SYN Flooding 攻击时，一旦防火墙承受不了攻击，整个内部网络都将瘫痪，而不是一台主机。可见，在这种情况下，使用防火墙的防御系统更弄巧成拙，说明基于防火墙和路由器的防御系统也不能有效地防御高强度的 SYN Flooding 攻击。应用层代理服务器本身会受到 SYN Flooding 攻击，不能起到保护作用。网关包过滤主要是指以屏蔽路由器对出入包进行过滤等。在 RFC2267[19]，Ferguson 和 Senie 提出了通过网络入口过滤来防止攻击者伪造 IP 地址发动 DoS 攻击。其优点是过滤速度快，对网络性能的影响小，费用低，但由于网关的包过滤机制目前仅是一种附带功能，过滤策略显得简单，最大的缺点还在于网关上的过滤策略因复杂性和安全性原因很难被用户动态改变。

另一类解决方案则基于服务器本身，包括 SYN cache[20] 和 SYN cookie[21] 技术。SYN cache[20] 对 TCP 协议的实现稍加修改，用哈希表代替线性表来保存半连接信息，减少了 TCP 待连接队列的存储空间开销。SYN cache 仅仅是部分改善了服务器抗 SYN Flooding 攻击的能力。SYN cookie[21] 则根本不维护任何半连接信息，它要求 TCP 软件对收到的 SYN 包应答特殊的 SYN/ACK 包，其应答号由源和目的地址、源和目的端口、连接发起序列号按加密算法求出。当客户方对此 SYN/ACK 包应答以 ACK 包时，服务器检查其应答号以判断是否为对先前 SYN/ACK 包的确认，如正确则直接进入连接建立状态，从而跳过半开连接状态，避免 SYN Flooding 攻击。SYN cookie 的缺点在于 TCP 协商选项被全部丢弃，从而对 TCP 性能有不可忽视的影响。另外，由于 SYN cookie 在连接建立起来之前根本不保存相关状态，因而 TCP 超时重发/出错重传的特点被摒弃，这直接影响到 TCP 面向连接的可靠传输特性。

上述几种解决方案，主要是防止固定网络中 SYN Flooding 攻击。由于在无线自组织网络中的泛洪攻击方法与 SYN Flooding 攻击不相同，Ad hoc Flooding 攻击不需要建立 TCP 连接，所以上述方案无法用于阻止无线自组织网络中的泛洪攻击。

对无线自组织网络中 DoS 攻击及其阻止方法的研究国际上才刚刚开始，相关论文还比较少见。Vikram Gupta 集中研究了无线自组织网络 MAC 层上 DoS 攻击的方式[3]。攻击者通过不公平的长期占用 MAC 层信道，导致其他节点无法得到通信信道进行正常报文交换。其主要原因是 MAC 层通信协议无法公平分配信道。I. Aad 对两种 DoS 攻击模型进行

了分析[4]：其一是 JellyFish 攻击模型，主要攻击 TCP 协议，其方法是打乱报文传送顺序、同期性抛弃报文和延迟转发报文。通过这三种方法的使用，引发 TCP 端到端的拥塞控制机制，从而导致端到端的流量大幅度降低甚至为 0；其二是黑洞攻击模型，攻击者参与网络路由查询，建立通过攻击者的路由，然后抛弃所有通过其节点的数据报文。Yih-Chun Hu 研究了 Rushing 攻击及其防御方式[14]，提出了通过随机转发的方法来对付泛洪攻击。其方法是将节点只处理第一个收到的路由查询报文，抛弃随后到达的相同路由查询报文，改为节点收集一定数量的相同路由查询报文，然后选择其中任意一个进行处理。这样就可以阻止泛洪攻击。Yih-Chun Hu 在另一篇论文[12]中研究了 Wormhole 攻击及其防御方式。提出了基于时间约束和基于空间约束的两种防御方法。基于时间约束的方法是给每个发出报文打下时间戳，接收时对比一下时间，将报文发送与接收控制在一定时间内。基于空间约束的方法是利用 GPS 信息，报文发送时带上位置信息，接收时可查看其发送距离，防止传送过远，将报文发送与接收控制在一定范围之内。

4.4 Ad hoc Flooding 攻击模型

本节提出一种新的攻击模型——Ad hoc Flooding 攻击。它能针对无线自组织网络中的所有采用按需路由协议发动 DoS 攻击，例如：DSR[8]、AODV[22]、LAR[23]等，甚至有些路由安全协议也不能幸免，如：SRP[24]、Ariadne[9]、ARAN[25]、SAODV[26,27]，因为它们只是提供节点相互认证，防止恶意节点修改路由协议报文其目的是防范外界的攻击，而对内部节点发动的 DoS 攻击丝毫不能防止，其安全认证的过程需要大量的计算，反而更增强 DoS 攻击的效果。下面基于 AODV 来描述 Ad hoc Flooding 攻击方法，针对其他路由协议的攻击方法类似。下面我们先概要介绍 AODV 路由协议，然后分别讨论攻击方式。

4.4.1 AODV 路由协议概述

在 AODV 路由协议中，节点之间的路由建立是完全按需进行的。当源节点需要发送报文至某个节点，而源节点中没有至该节点的路径时，源节点可广播路由请求报文(RREQ)来查询到目标节点的路径。RREQ 报文在整个网络中泛洪发送。节点收到 RREQ 时，首先检查是否是以前接收过的 RREQ 报文，若是就直接抛弃不进行处理，若不是进行下一步处理。收到 RREQ 报文的节点，搜索本节点路由表中是否记录有到目的节点的路由。若没有，节点暂存 RREQ 报文的源节点地址、目的节点地址、上游节点地址和目的序列号，建立反向路由后，将 RREQ 重新广播出去。若有到目标节点的路由或本机就是目的节点，可发送“路由回答”(RREP)沿 RREQ 来时建立的反向路由回到源节点，回答中包含更新的目的序列号，沿途转发“路由回答”的中间节点根据回答更新本机路由表，设置路由的下游节点、目的序列号、有效时间信息。当 RREP 到达源节点时，就建立了一条源节点到目标节点的路由。

AODV 通过周期性的广播 hello 报文来监视链路状态，若节点在使用某个链路发送报文时发现该链路中断，节点将从路由表中删除包含该断开链路的路由，并发送“路由出错”报文(RRER)至路由起始节点，通知那些因链路断开而不可达的节点将对应路由从路由表中删除，沿途转发 RRER 的节点也删除自己路由表中的对应路由。当 RRER 到达源节点时，源节点也删除这条路由，如果还需要发送数据，源节点可重新进行路由查询。

4.4.2 Ad hoc Flooding 攻击方法

在 AODV 路由协议中,泛洪查找路由是非常消耗网络资源的,为了减少泛洪 RREQ 报文对网络的影响,AODV 协议采取了一些措施。首先设置了 RREQ 每秒最大发送数,每个节点在一秒内发送的 RREQ 报文数不能超过这个数值。其次,节点在发送的 RREQ 报文后,要设置一个最大查询往返时间,等候 RREP 的返回,如果超过最大查询往返时间没有收到节点回答才能准备重新的发送 RREQ 报文,但也不能立即发送,需要等待一段时间,该时间长短为 RREQ 查询往返时间的两倍。再次,RREQ 的泛洪查询范围必须依次递增,通过 RREQ 报文中的 TTL(time-to-live)进行控制,开始时设置范围小,查询不到时,再依次增加,直至收到 RREP 或达到最大限制。AODV 路由协议通过上述方法来控制泛洪 RREQ 查找的频率与范围,减少对网络资源的消耗。

但在 Ad hoc Flooding 攻击中,入侵者不顾这些规定,尽力消耗网络资源,攻击分为两步。第一步,入侵者选择路由查询的节点地址。如果它知道整个网络的地址范围,它将选择不在网络内 IP 地址作为路由查询的节点地址,因为没有节点能够回答它的 RREQ 报文,每个节点就要一直暂存的 RREQ 的信息和反向路由,直至超时才能删除这些信息,能够尽可能长时间占用资源。如果入侵者不知道整个网络的地址范围,它就随机选择一些 IP 地址进行路由查询。第二步,入侵者以选择好的 IP 地址为目标,大量、连续地发送 RREQ 报文。不管 AODV 设置的 RREQ 每秒最大发送数,尽力多发送 RREQ,同时直接将 TTL 设置为最大值,在全网内泛洪查找。如果发送 RREQ 的地址用完,就开始新一轮的发送,不顾 RREQ 的查询往返时间和退避时间。当入侵者采用上述方法发动 RREQ Flooding 攻击时,整个网络就会充满 RREQ 报文,导致通信带宽和节点两方面的资源枯竭。连续不断的 RREQ 在网络中泛洪发送,占用了大量无线通信带宽,导致网络拥塞,正常通信无法进行。对于节点来说,每收到一个 RREQ 报文,从上节 AODV 协议概述可知,它都要缓存 RREQ 报文的源节点地址、目的节点地址、上游节点地址和目的序列号并建立反向路由,该缓存要等待 RREP 或超时后才能释放。如果没有 RREP 到达,又不断接收新的 RREQ 报文,有限的缓存就会被消耗完毕。此时,如果其他节点要建立路由,再发送 RREQ 报文,这些节点就不能接收新的 RREQ 报文,导致正常的路由建立无法进行。图 4-4 显示一个 RREQ flooding 攻击的例子。攻击节点 H 向周围节点泛洪发送攻击报文,周围节点收到后继续泛

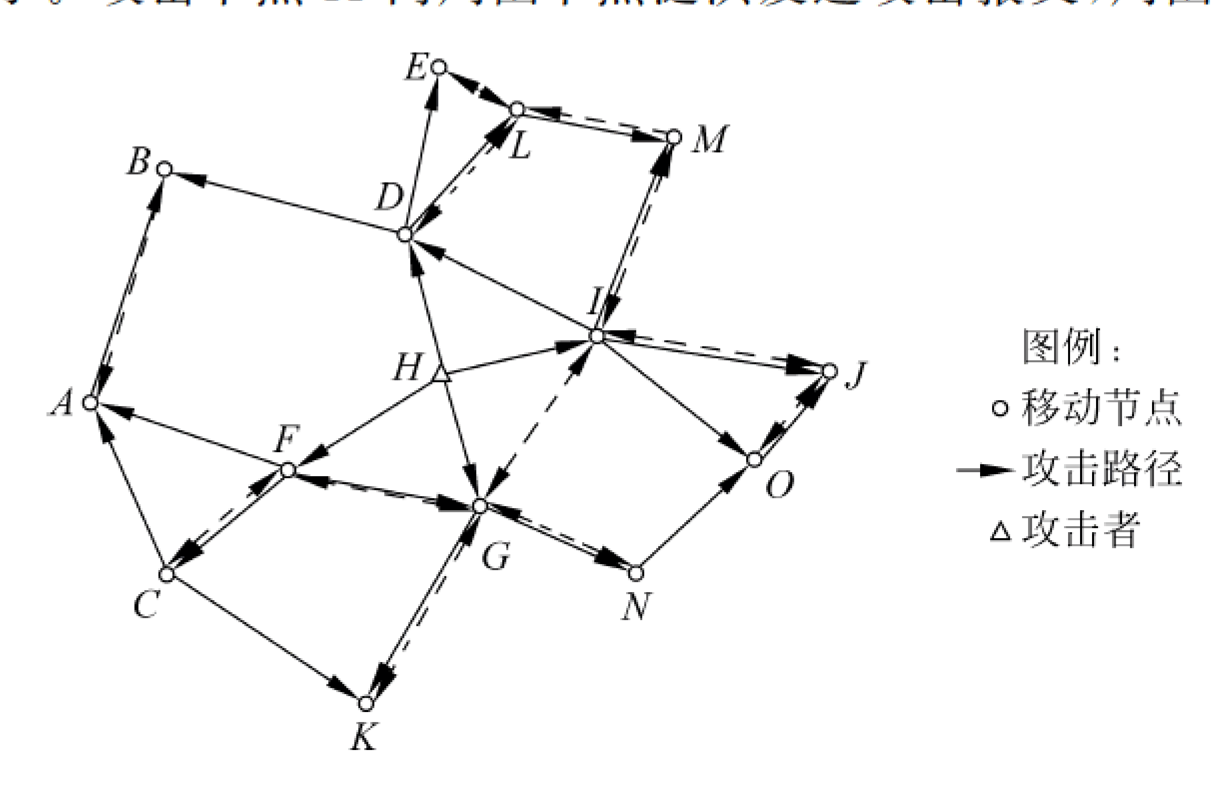


图 4-4　Ad hoc Flooding 攻击示意图

洪传播，造成整个网络充满了攻击报文，网络性能严重下降。

4.4.3 Ad hoc Flooding 攻击与 SYN Flooding 攻击的异同

不同于上述 Ad hoc Flooding 攻击方法，SYN Flooding 攻击方法为：攻击者通过伪造的 IP 地址发送大量的 TCP 连接请求至某个节点，每个 TCP 请求都会导致目标节点建立一个连接缓冲。该连接缓冲由于连接无法建立而不能释放，一旦节点缓冲空间用尽，就不能接收新的 TCP 连接请求，从而导致了拒绝服务。SYN Flooding 攻击是通过大量发送假的 TCP 连接请求的方法，造成某个被攻击机器因资源耗尽而瘫痪。

表 4-1 从攻击方法、目标、结果等方面对两种不同的攻击方法进行了对比。Ad hoc Flooding 攻击与 SYN Flooding 攻击的目的是类似的，都是导致拒绝服务。但 Ad hoc Flooding 攻击通过泛洪发送大量攻击报文的方式，导致整个网络性能下降，不能正常通信。SYN Flooding 攻击通过发送大量假的 TCP 连接到某个主机，导致单个主机不能正常工作。上述分析表明 Ad hoc Flooding 攻击与 SYN Flooding 攻击是两种不同的 DoS 攻击模型。

表 4-1　Ad hoc Flooding 攻击与 SYN Flooding 攻击的对比

攻 击 名 称	SYN Flooding 攻击	Ad hoc Flooding 攻击
攻击方法	大量假的 TCP 连接	泛洪发送过量路由请求报文
攻击目标	网络中单个主机	整个无线自组织网络
受攻击的协议	TCP 协议	按需路由协议
攻击所在的网络协议层次	传输层	网络层
攻击结果	单个主机瘫痪	整个网络性能下降

4.5　防止 Ad hoc Flooding 攻击的方法

本节提出通过邻居阻止的方法，来防止 Ad hoc Flooding 攻击的方式。在论述这种方法之前，我们设定一些前提条件：链路层通信是双向的；无线通信是无方向性的，任何一个节点发送报文其周围节点均能收到。节点 MAC 地址与 IP 地址一一对应，且不可任意改变。

无线自组织网络的其中一个特点是多跳的通信，即节点之间的通信必须通邻居节点和中间节点进行转发，如果邻居节点拒绝转发某个节点报文，则该节点就会被隔绝于网络。图 4-5 是一个无线自组织网络的例子，网络中有 15 个节点，其中 H 是攻击者。节点 H 要与其他节点之间的通信必须通过节点 D、F、G、I 进行中转。如果节点 H 的周围邻居节点 D、F、G、I 都拒绝接收节点 H 的报文，并中断与节点 H 的通信链路，节点 H 就会被隔绝于网络，不能再危害网络。

根据上述思想来设计防御方法。在 AODV 路由协议中，节点接收处理 RREQ 报文是依据先来先处理的原则进行的，如图 4-5 所示，如果节点 F 先收到攻击者 H 发送的大量 RREQ 报文，此时如果节点 A 再发送 RREQ 报文，节点 F 要先处理完节点 H 的报文后才能处理节点 A 的报文。如果节点 H 发送大量的 RREQ 报文导致节点 F 路由表缓存溢出，那么节点 F 就无法接收节点 A 的 RREQ 报文，这就导致了节点 F 的拒绝服务。为了阻止这种攻击，将这种先来先处理的原则改为基于优先级的处理原则，节点为每个邻居节点设立

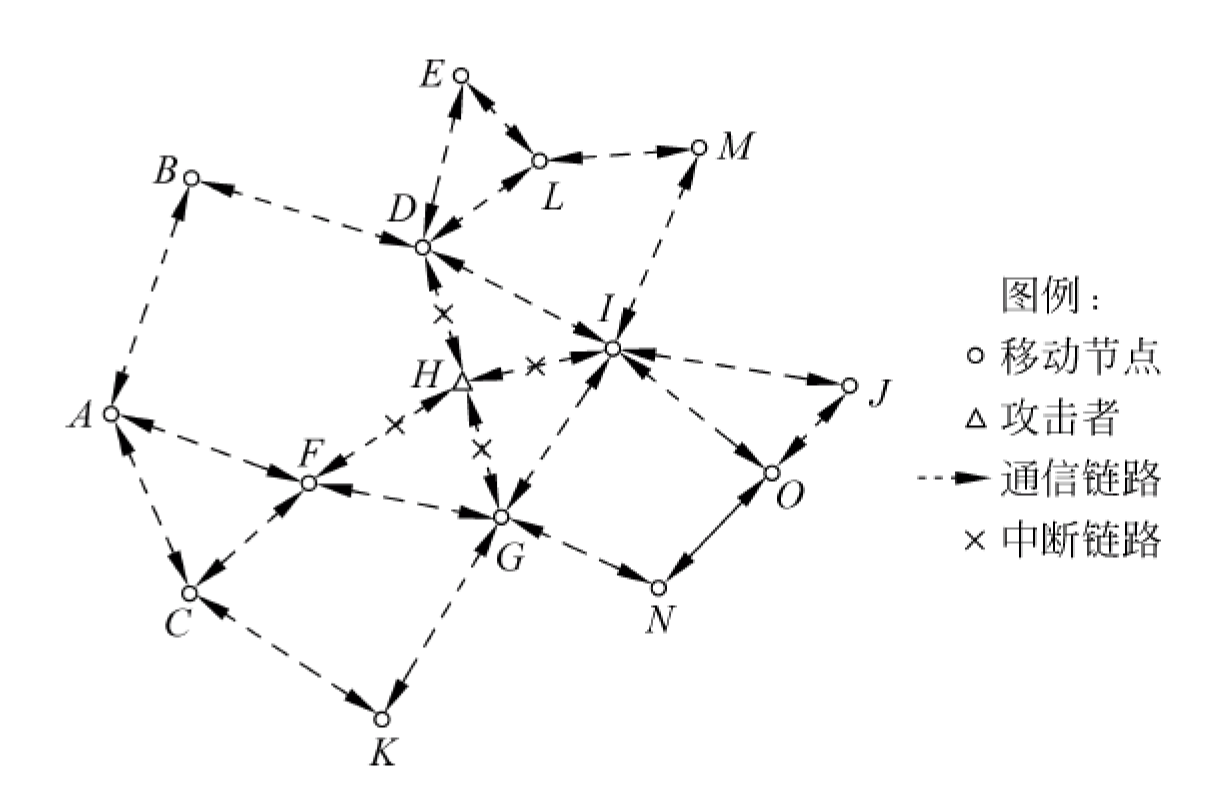


图 4-5 邻居节点隔绝攻击者

一个处理优先级，其优先级的大小与该节点以前的发送频率有关，节点处理报文时先查优先级，优先级高的节点发送的报文首先处理。例如，初始值都可设为 1，当节点 *H* 前 1 秒发送了 10 个 RREQ 报文，那么其周围邻居节点如 *F* 就将节点 *H* 的优先级改为 1/10。这时如果节点 *A* 发送 RREQ 报文，如果在节点 *F* 中节点 *A* 的优先级为 1，节点 *F* 就先处理节点 *A* 的 RREQ 报文，而将节点 *H* 发送的 RREQ 报文排在后面处理。为进一步防止过量发送 RREQ 报文，将 AODV 协议中规定的节点每秒最大发送 RREQ 报文数设立为阈值，当某个节点的发送频率超过这个值时，就认定为该节点为攻击者，此后拒绝接收该节点的报文。采用这种方法以后，攻击者 *H* 如果发送大量的 RREQ 报文，其周围节点就会迅速降低对节点 *H* 的处理优先级，如果发送 RREQ 频率超过阈值，邻居节点就会拒绝接收该节点的报文，从而阻止了 RREQ Flooding 攻击行为。

4.6 模拟实验

4.6.1 实验设置

为了研究 Ad hoc Flooding 攻击对网络性能产生的影响，同时验证防御方法的实际作用效果，我们进行了一系列网络模拟实验。实验分为两类，第一类是在网络模拟器上 AODV 协议中实现了 Ad hoc Flooding 攻击，模拟在不同泛洪攻击强度下网络性能的变化。第二类在 AODV 协议中实现了邻居阻止的防御方法，进行了多个场景的实验，模拟网络在增加了防御方法后，网络性能的变化情况。

实验平台为 Pentium 4，配置为 CPU 主频 1.8GHz，RAM 容量 512MB，使用的操作系统是 Red Hat Linux 7.2，网络仿真平台是 NS-2 2.26(Network Simulator Version 2.26)[28]。它是美国国防高级研究计划局资助的 VINT(Virtual InterNet Testbed)[29]项目开发的专用于网络模拟实验的软件工具。NS-2 是一个开放源码免费软件，带有大量协议库的支持。主要在学术界用于对网络协议的算法验证。初始时主要针对固定网络的模拟，CMU(卡内基·梅隆大学)对 NS-2 进行了扩展，在物理层、链路层、MAC 层等方面增加了对于无线网络的支持，用这些增加的部件可以对无线子网、无线局域网、无线自组织网络、移动 IP 等进行仿真。

移动节点采用 IEEE 802.11 DCF 作为 MAC 层的协议,RTS/CTS 控制报文用在向邻居节点发送单播数据报文过程中,实施“虚载波侦听”和信道预留来减少“隐藏终端”的影响。数据报文发送后有一个 ACK 报文。RTS 和广播报文使用物理载波侦听来发送,采用 CSMA/CA 来发送这些数据报。节点电磁波的辐射范围为 250m。节点间通信信道带宽为 2Mb/s。

本模拟实验节点运动范围为 1000m×1000m 的方形区域,其中分布 50 个移动节点。节点移动选用常用的 random waypoint 方式。在指定的范围内,节点随机选择某个目标后,在预先设点的最大速率和最小速率中随机选取一个速率匀速前进,当节点到达该目标时则停留预设的时间。利用速率和停留时间这两个参数,可以唯一确定整个网络的移动模型。模拟时间为 900s。模拟实验场景如图 4-6 所示,其中设置一个攻击者,为 30 号移动节点。

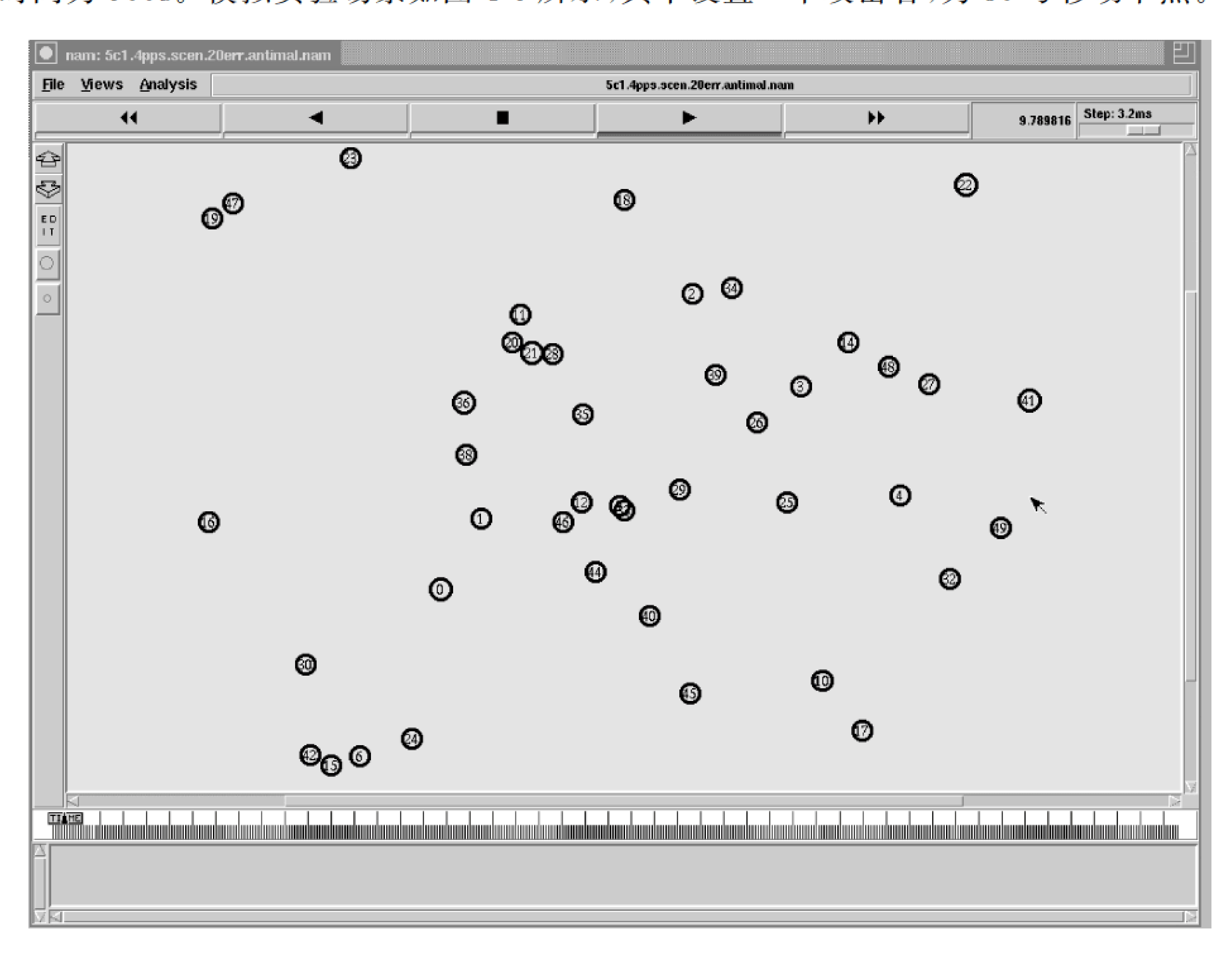


图 4-6　模拟实验场景

测试收集两种数据,分组传递率(packet delivery rate):应用层信源发送的分组数目与信宿接收分组数目之比。它描述的是通过应用层观察到的报文丢失率,又反映了网络所支持的最大吞吐量。它是路由协议完成性和正确性的指标。平均延迟(average delay):它是报文从源节点到目标节点的平均传输时间,它反映了网络传输性能。

4.6.2　Ad hoc Flooding 攻击实验结果

为了测试在不同攻击强度下网络性能所受的影响,我们进行了五组实验。第一组是没有 Ad hoc Flooding 攻击的情况下,网络的报文传递率和报文传送平均延迟。第二组是攻

击者每秒发送 10 个泛洪攻击报文的情况下,网络的报文传递率和报文传送平均延迟。第三组是攻击者每秒发送 20 个泛洪攻击报文。第四组攻击者每秒发送 30 个泛洪攻击报文。第五组攻击者每秒发送 40 个泛洪攻击报文。模拟实验时间 900s,每 100s 统计一次网络的报文传递率和报文传送平均延迟。实验运行到 100s 时,统计 0～100 秒之间的网络报文传递率和报文传送平均延迟。实验运行到 200s 时,统计 100～200 秒之间的网络报文传递率和报文传送平均延迟。其余的依此类推,实验运行期间共统计 9 次。程序设定 0～300 秒之间,节点正常运行,300s 以后攻击者开始发动 Ad hoc Flooding 攻击。图 4-7 显示 30 号攻击者从 300s 开始发动第一波泛洪攻击。

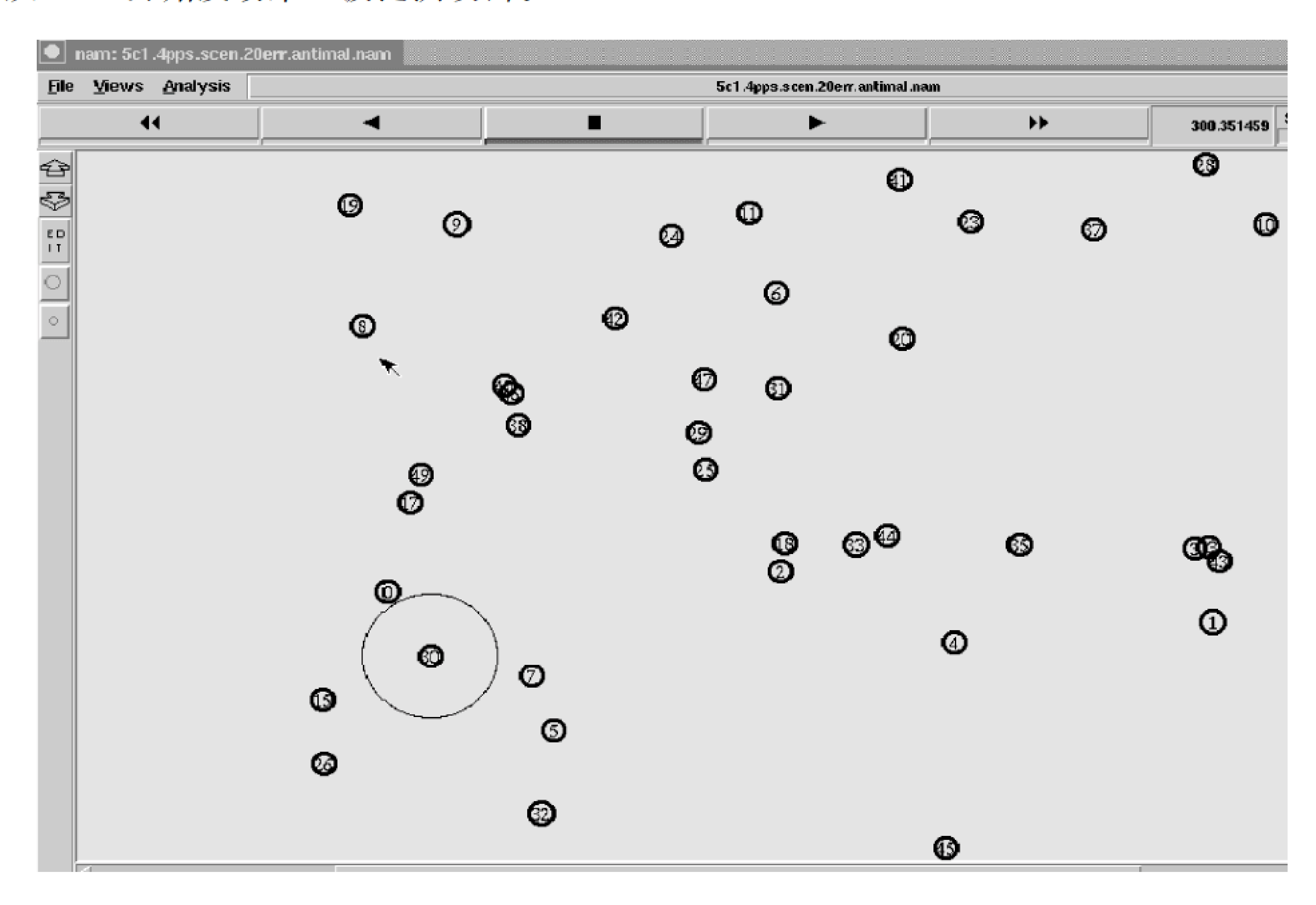


图 4-7 Ad hoc Flooding 攻击开始

图 4-8 显示当网络中节点收到第一次泛洪攻击报文后,各节点再进行转发。攻击者连续发送攻击报文,每个节点收到攻击报文后,又要进行转发,如此进行下去,整个网络的通信带宽将被攻击者完全消耗。图 4-9 显示网络中充满了攻击者发送的攻击报文。

对上述五种情况下的报文传递率按每 100s 的间隔进行了汇总统计,计算在每 100s 内的五个发送源平均报文传递率。图 4-10 显示五个场景下的模拟实验结果,每条实验曲线代表一种攻击强度下报文平均传递率,第一条曲线在最上面代表没有攻击的情况下,第二条曲线在第一条曲线的下面,为 10 个攻击报文的情况下,以此类推,第五条曲线在最下面,为 40 个攻击报文情况下的报文平均传递率。从图 4-10 中可以明显看出,当网络中没有攻击的情况下,第一条曲线显示报文平均传递率接近 100%。当网络中出现攻击报文时,传递率开始下降。当攻击强度为每秒 10～20 个泛洪攻击报文时,网络性能下降不明显,第二、三条曲线显示报文平均传递率保持在 80%左右,也就是说大部分发出的报文能够被传送到目的节点。其主要原因是攻击报文消耗的带宽还不太多,网络还能够保持正常运行,甚至还有一定程度的恢复。但是当 Ad hoc Flooding 攻击强度达到每秒 30 个泛洪攻击报文时,网络已经

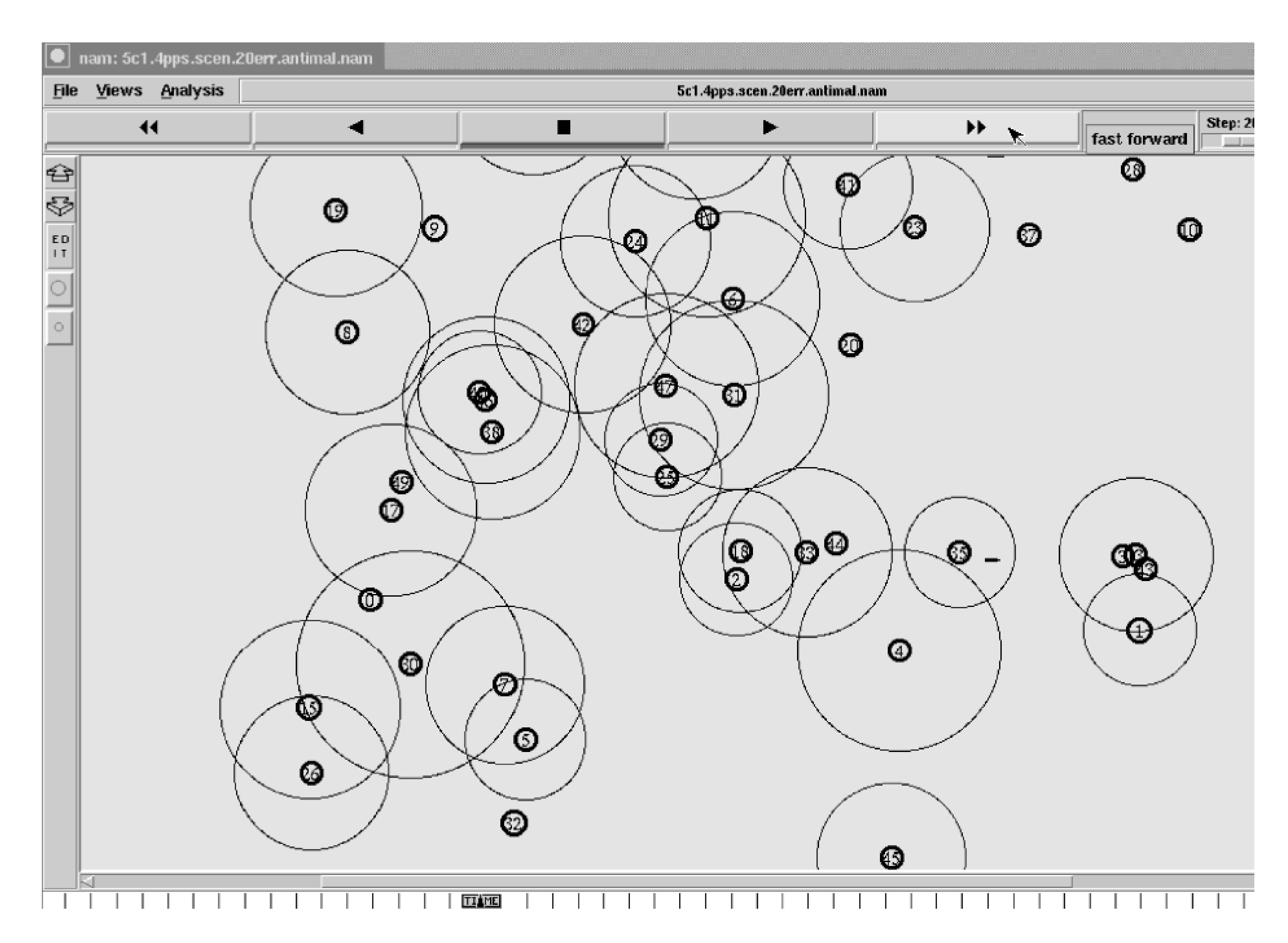


图 4-8　节点收到泛洪攻击报文后开始转发

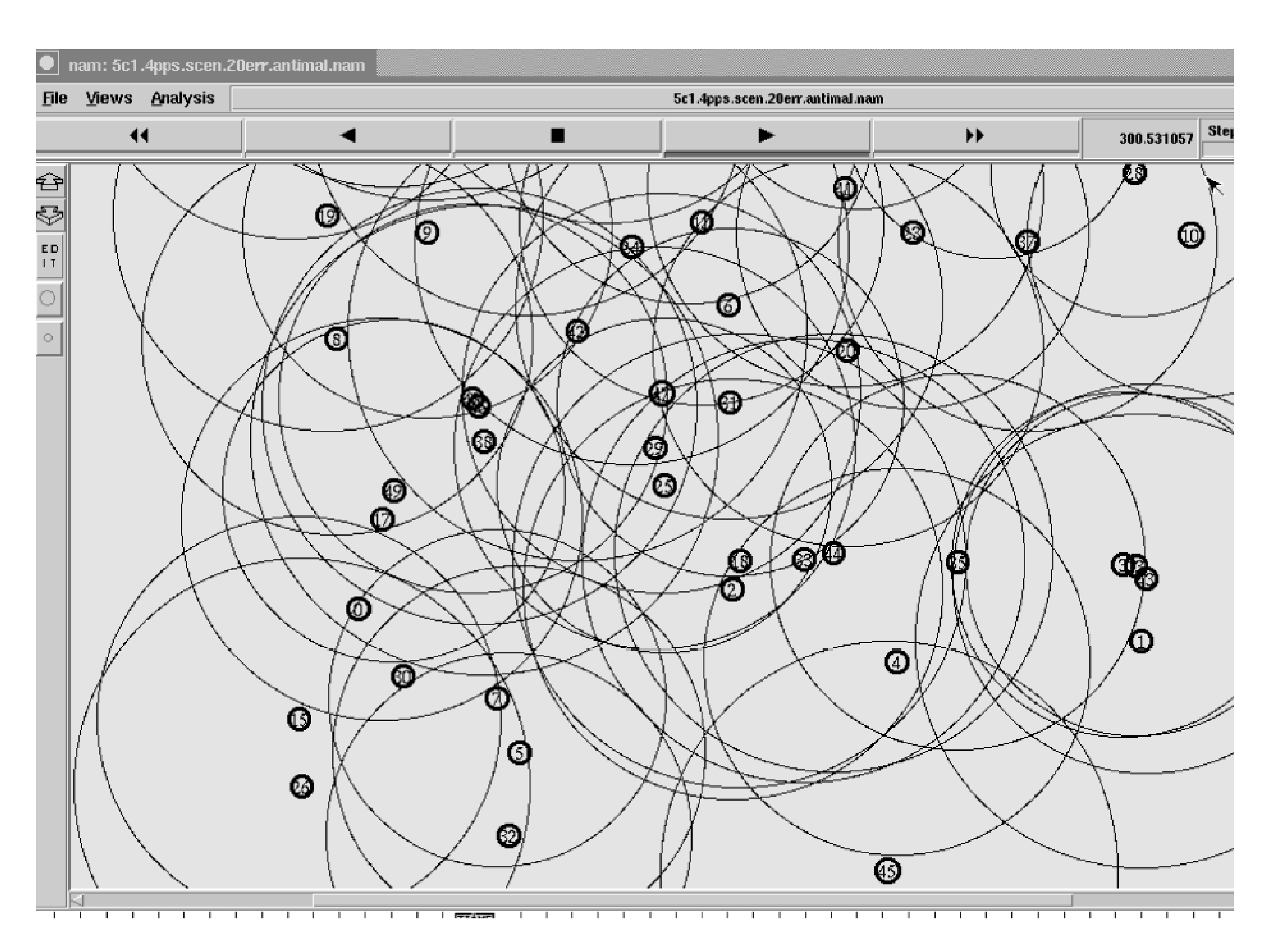


图 4-9　网络中充满了攻击报文

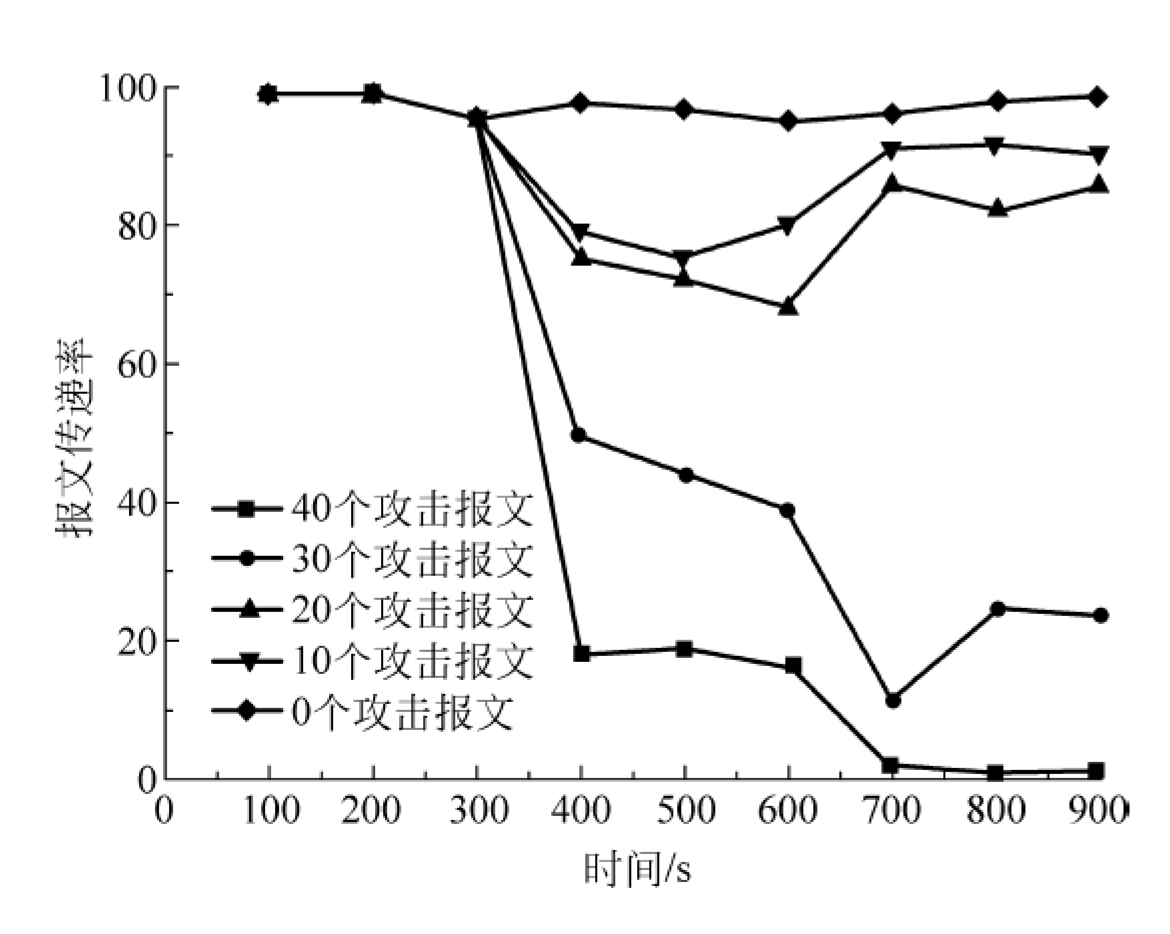


图 4-10　在 Ad hoc Flooding 攻击下的报文传递率

不能保持正常运行,性能明显下降,第四条曲线显示,报文平均传递率下降到 30%左右,也就是有 2/3 的报文由于网络拥塞被抛弃了。当 Ad hoc Flooding 攻击强度达到每秒 40 个泛洪攻击报文时,如第五条曲线显示,网络性能急剧下降,报文平均传递率几乎降到了零,也就是说整个网络除了泛洪攻击报文,其他报文都被抛弃了。

4.6.3　Ad hoc Flooding 防御方法实验结果

为了验证所提出的防御方法的有效性,我们对按需路由协议 AODV 进行了更改,将所提出的防御方法加入到节点对路由报文的处理之中。进行了六个场景下的实验,第一个场景为没有攻击的情况下,第二个场景为攻击强度为每秒 10 个泛洪攻击报文的情况下,依此类推,第六个场景为每秒 50 个泛洪攻击报文的情况下。其实验结果如下。

没有攻击的情况下,从图 4-11 可以看出,平均传递率达到了 97%。Ad hoc Flooding 攻击时,为了对比没有攻击、存在攻击、存在防御三种情况下的性能,我们专门设计了如下场景:前 300s 没有攻击,从 300s 以后启动 Ad hoc Flooding 攻击,从 600s 开始防御方法开始发挥作用,也就是说,300～600s 只存在攻击,600～900s 攻击和防御同时存在。与上节同样,每 100s 计算一次报文平均传递率。图 4-12 显示在每秒 10 个攻击包的情况下的报文平

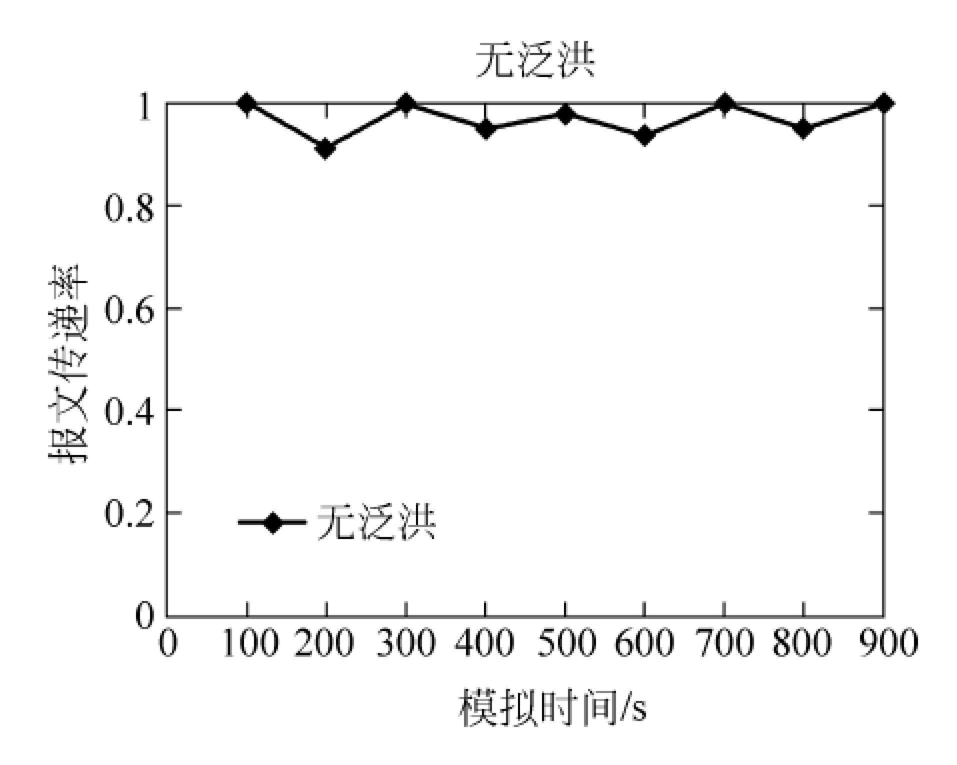


图 4-11　没有攻击时的报文传递率

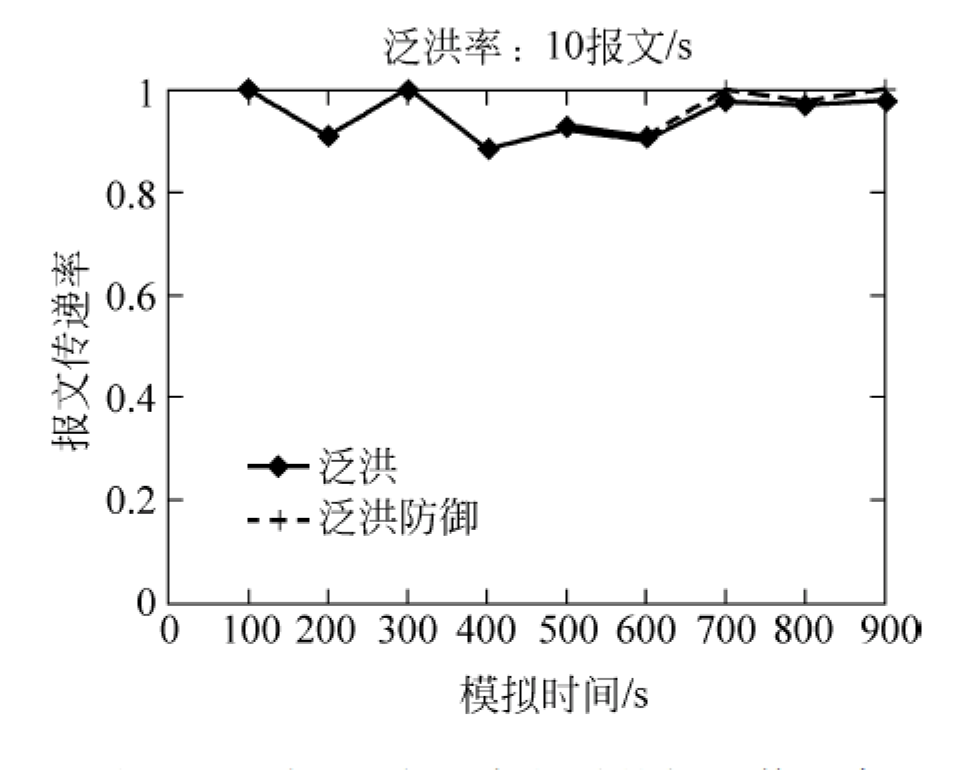


图 4-12　在 10 个攻击包时的报文传递率

均传递率，可以看出 300s 后攻击时性能有所下降，增加防御后性能有所提高，图中虚线表示启动防御后的报文传递率，但不明显，因为攻击所造成的性能下降不明显。

从图 4-13 和图 4-14 看出，当攻击包为每秒 20 个时，防御效果开始显现，仍然不够明显。但当攻击包为 30 每秒时，网络性能下降非常明显，300s 开始攻击时，报文传递率下降到 50％左右，当启动防御后的报文传递率开始恢复，从 50％升到 80％。启动防御后，阻断了攻击包对整个网络的影响，但攻击者仍然会对其周围节点产生影响，同时防御算法也会产生一些消耗，所以不能完全将网络性能恢复到没有攻击的情况下。

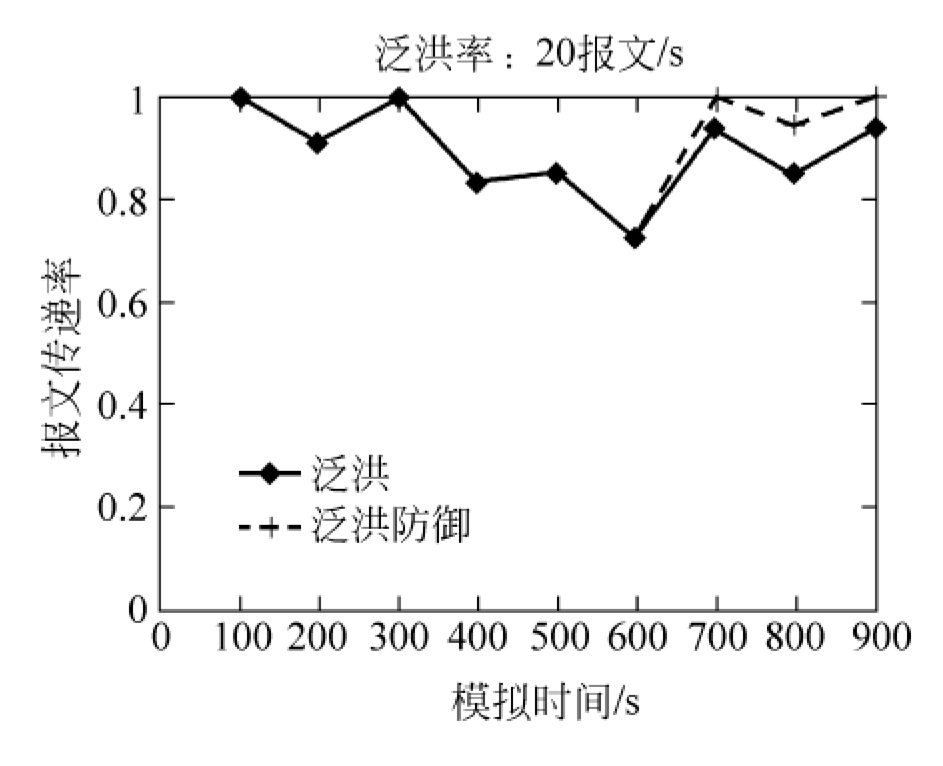


图 4-13　在 20 个攻击包时的报文传递率

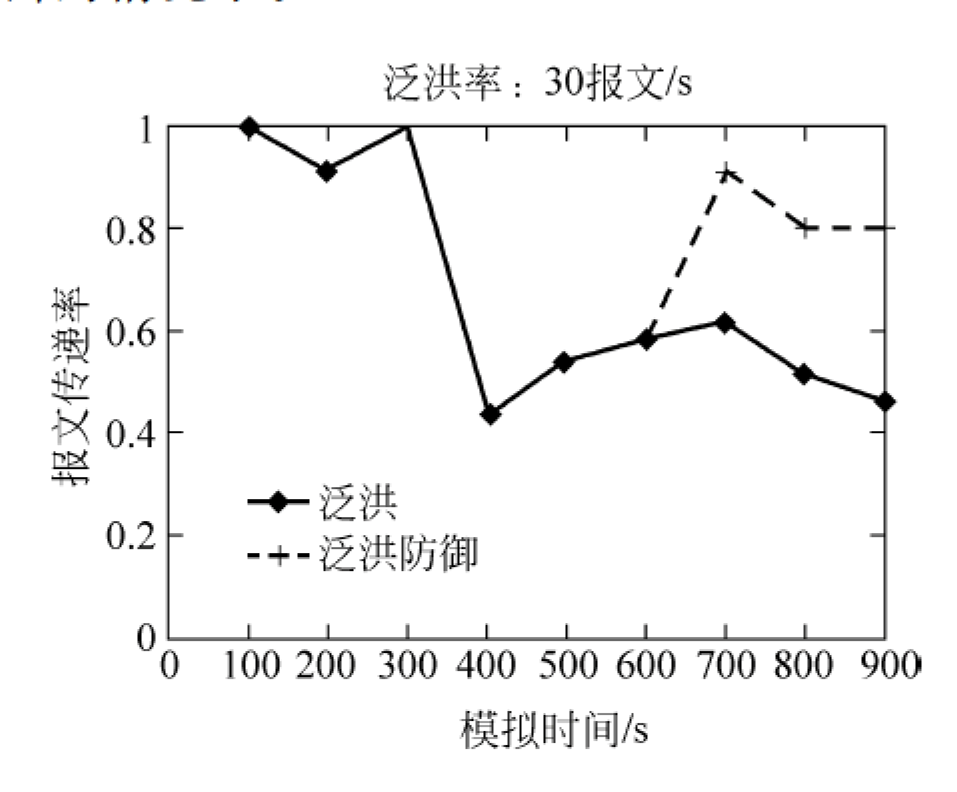


图 4-14　在 30 个攻击包时的报文传递率

当 Ad hoc Flooding 攻击强度达到每秒 40 个以上时，网络报文传递率就会降到 20％以下，从图 4-15 和图 4-16 就可以看出。但当 600s 以后，启动防御后，网络性能明显提高，报文传递率上升到 80％。从上图中还可以发现，无论 Ad hoc Flooding 强度增加多少，防御都能将报文传递率恢复到 80％，说明防御方法能够成功地阻止攻击，没有漏掉一个攻击包。

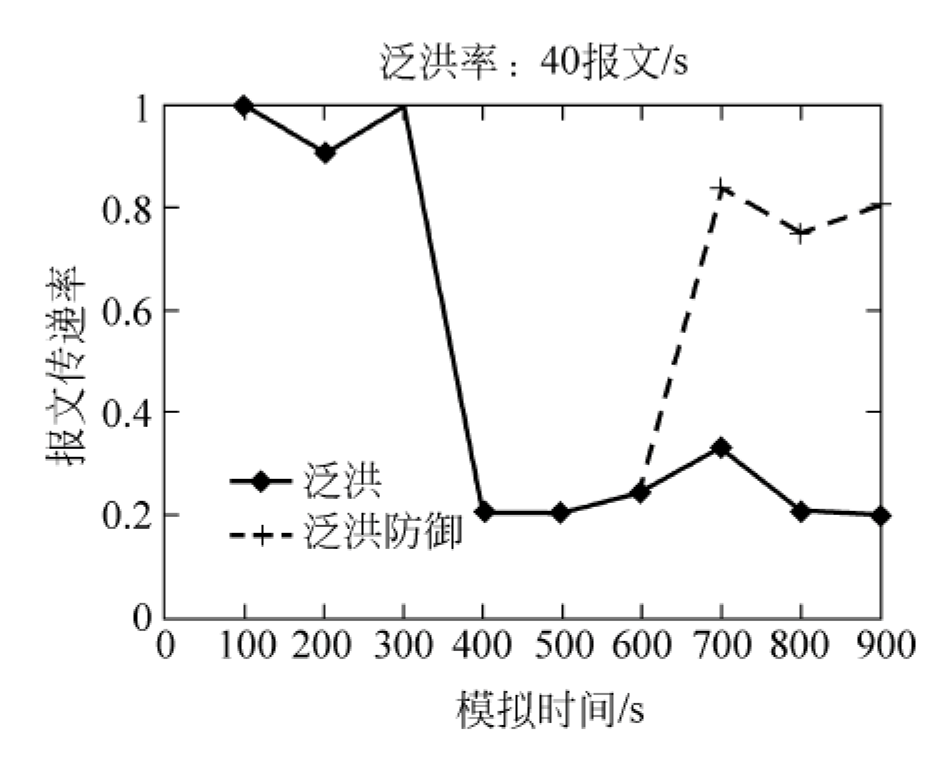


图 4-15　在 40 个攻击包时的报文传递率

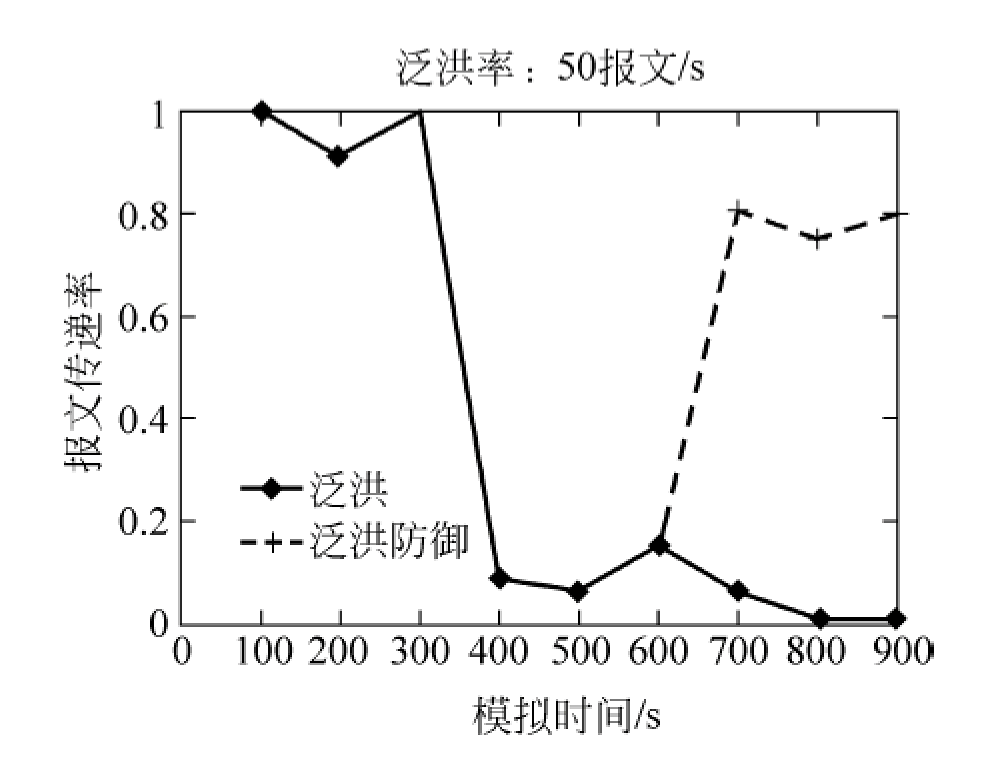


图 4-16　在 50 个攻击包时的报文传递率

图 4-17 显示的是上述六种场景下的平均报文传输延迟，同样，每 100s 统计一次，可以发现在有攻击情况下，报文传输延迟随攻击强度增加而增加，启动防御后，传输延迟明显回落，几乎接近没有攻击时的情况。

从上述实验结果可以得出，当 AODV 协议中增加了对 Ad hoc Flooding 攻击的防御方法后，明显提高网络吞吐量和传输效率。

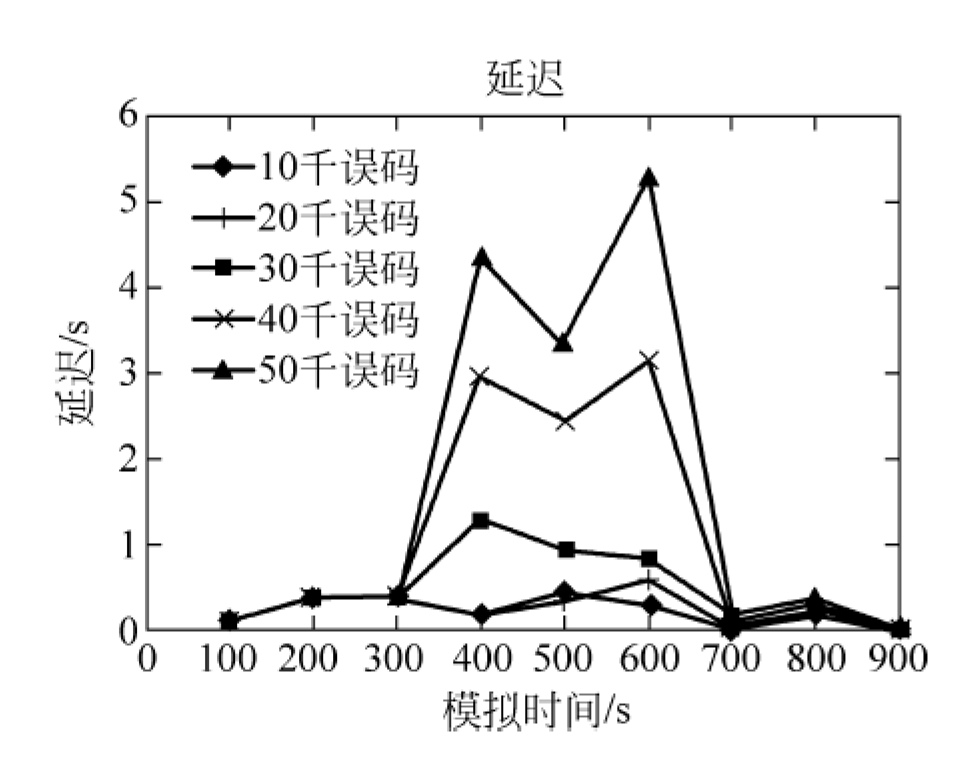


图 4-17 Ad hoc 泛洪攻击时的平均报文传输延迟

4.7 小结

本文提出一种新的 DoS 攻击模型——Ad hoc Flooding 攻击，该攻击主要针对于无线自组织网络中按需路由协议，通过 RREQ Flooding 和 DATA Flooding 两种攻击方法，即攻击者通过向网络泛洪发送过量的路由查询报文同时向所有节点发送攻击数据报文，消耗网络带宽和节点资源，导致网络拥塞和节点无法正常通信，造成网络性能严重下降。为了阻止 Ad hoc Flooding 攻击，提出了两种防御策略：邻居阻止和路径删除。模拟实验证实，通过这两种方法的结合，能够成功地发现并阻止 Ad hoc Flooding 攻击，提高了网络性能。

参考文献

[1] Kargl F, Maier J, Weber M. Protecting web servers from distributed denial of service attacks. In Proceedings of 10th International World Wide Web Conference, May 2001.

[2] Mirkovic J, Reiher P. A taxonomy of DDoS attack and DDoS defense mechanisms. ACM SIGCOMM Computer Communication Review, 2004, 34(2): 39-53.

[3] Vikram Gupta. Denial of Service Attacks at the MAC Layer in Wireless Ad hoc Networks. Proceedings of MILCOM 2002, 2(7-10): 1118-1123.

[4] Aad I, Hubaux J P, Knightly E. Denial of Service Resilience in Ad hoc Networks, Proceedings of the 10th annual international conference on Mobile computing and networking (MobiCom2004), Philadelphia, October 2004.

[5] Corson S, Macker J. Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations, RFC 2501, January 1999.

[6] Hubaux J P, Levente Buttyan. Srdjan Capkun The Quest for Security in Mobile Ad hoc Networks Proceedings of the 2001 ACM International Symposium on Mobile Ad hoc Networking & Computing 2001, Long Beach, CA, USA.

[7] Zhou Lidong, Zygmunt J Haas. Securing Ad hoc Networks. IEEE Networks Special Issue on Network Security, November/December, 1999.

[8] Johnson D B, Maltz D A, Hu Y C. The Dynamic Source Routing Protocol for Mobile Ad hoc Networks (DSR), Internet-Draft, draft-ietf-manet-dsr-09. txt, 15 April 2003, http://www. ietf. org/internet-drafts/draft-ietf-manet-dsr-09. txt.

[9] Kimaya Sanzgiri, Bridget Dahill, Brian Neil Levine, et al. A Secure Routing Protocol for Ad hoc Networks, Proc of 2002 IEEE International Conference on Network Protocols (ICNP), Paris, 2002: 78-89.

[10] Frank Stajano, Ross Anderson. The Resurrecting Duckling: Security Issues for Ad-hoc Wireless Networks, The 7th International Workshop on Security Protocols, LNCS 1796, Springer-Verlag, Cambridge, United. Kingdom, 1999: 172-194.

[11] Hu Y C, Perrig A, Johnso D B. Wormhole Detection in Wireless Ad hoc Networks, Technical Report TR01-384, Department of Computer Science, Rice University, December 2001.

[12] Hu Y C, Perrig A, Johnso D B. Packet Leashes: A Defense against Wormhole Attacks in Wireless Networks, Proc of the Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies(INFOCOM 2003), San Francisco, April, 2003: 1976-1986.

[13] Deng Hongmei, Li Wei, Dharma P. Agrawal, Routing Security in wireless Ad hoc Networks, IEEE Communications Magazine, October 2002: 70-75.

[14] Hu Y C, Adrian Perrig, Johnson D. Rushing Attacks and Defense in Wireless Ad hoc Networks Routing Protocols, In Proceedings of the ACM Workshop on Wireless Security (WiSe 2003), September 19 2003, Westin Horton Plaza Hotel, San Diego. California.

[15] Chang K C. Defending against Flooding-Based Distributed. Denial-of-Service Attacks: A Tutorial, IEEE Communications Magazine, October 2002: 42-51.

[16] Schuba C, Krsul I, Kuhn M, et al. Analysis of a Denial of Service Attack on TCP, Proceedings of the 1997 IEEE Symposium on Security and Privacy.

[17] Wang Haining, Zhang Danlu, Kang G. Shin. Detecting SYN Flooding Attacks, IEEE INFOCOM' 2002, New York City, 2002.

[18] Karthik Lakshminarayanan, Daniel Adkins, Adrian Perrig, et al. Taming IP Packet Flooding Attacks. Computer Communication Review, 2004, 34(1): 45-50.

[19] Ferguson P, Senie D. Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing, RFC 2267, January 1998.

[20] Lemon J. Resisting SYN flood DoS attacks with a SYN cache, In Proceedings of USENIX BSDCon2002, San Francisco, California: 2002, 89-98.

[21] http://cr. yp. to/syncookies. html.

[22] Perkins C E, Belding-Royer E M, Das S R. Ad hoc On-Demand Distance Vector (AODV) Routing, RFC 3561, July 2003, http://www. ietf. org/rfc/rfc3561. txt.

[23] Ko Y B, Nitin Vaidya, Location-Aided Routing (LAR) in Mobile Ad hoc Networks. In Proceedings of the Fourth International Conference on Mobile Computing and Networking (MobiCom' 98), October 1998: 66-75.

[24] P. Papadimitratos, Z. Haas, Secure routing for mobile Ad hoc Networks, in Proceedings of the SCS communication Networks and Distributed Systems Modeling and Simulation Conference, San Antonio, TX, January 2002: 27-31.

[25] Hu Y C, Adrian Perrig, Johnson D B. Ariadne: A secure On-Demand Routing Protocol for Ad hoc Networks, in Proceedings of the MobiCom 2002, Atlanta, Georgia, September 2002: 23-28.

[26] Manel Guerrero Zapata. Secure Ad hoc On-Demand Distance Vector Routing. ACM Mobile Computing and Communications Review (MC2R), 2002, 6(3): 106-107.

[27] Manel Guerrero Zapata, Asokan N. Securing Ad-hoc Routing Protocols, In Proceedings of the 2002 ACM Workshop on Wireless Security (WiSe 2002), September 2002: 1-10.

[28] The Network Simulator ns-2, http://www. isi. edu/nsnam/ns/index. html.

[29] Virtual InterNetwork Testbed, http://www. isi. edu/nsnam/vint/index. html.

CHAPTER 5

第5章　无线自组织网络入侵检测研究

摘要：在无线自组织网络环境下，因为移动节点可能被攻击截获，从而泄露合法密钥，导致攻击从内部产生，传统的网络安全措施，如防火墙、加密、认证等技术，在无线自组织网络中难以应用。因此，只有通过入侵检测才能发现并清除入侵者。本文提出一种基于时间自动机分布式合作的入侵检测算法。首先，将整个网络分为各个监视区域，每一区域随机选出簇头担任监视节点，负责本区域的入侵检测。其次，按照DSR路由协议构筑节点正常行为和入侵行为的时间自动机，监视节点收集其邻居节点的行为信息，利用时间自动机分析节点的行为，发现入侵者。本算法不需要事先进行数据训练并能够实时检测入侵行为。最后，通过模拟实验证实了算法的有效性。

关键字：无线自组织网络、路由协议、网络安全、入侵检测、时间自动机。

5.1　引言

无线信道、动态拓扑、合作的路由算法、缺乏集中的监控等都使得无线自组织网络安全更加脆弱，特别是移动节点缺乏物理保护，容易被偷窃、捕获，落入敌手后重新加入网络，导致攻击从内部产生。而采用密码学理论的网络安全方案无法对抗此类攻击。此外，网络安全的发展史告诉我们没有100%的安全方案，无论多么安全的方案都可能存在这样或那样的漏洞。因此，入侵检测就理应成为安全方案之后的第二道防护墙。

为了保障无线自组织网络的安全，至今已经提出了许多安全解决方案[1]。但这些安全方案主要集中于密钥的分配与认证[2]、路由安全算法两个方面[3,4]。密钥的设置与认证和路由安全算法，这两种可以称为入侵阻止技术，所谓入侵阻止就是利用加密、认证、防火墙等技术来防止系统遭受外界的攻击。这些措施用于无线自组织网络之中，能够发挥一定的安全防范作用。但是，由于无线自组织网络中节点可任意移动，当网络处于敌对环境时，节点可能被截获而泄露密钥，敌方节点可持密钥冒充合法节点加入网络进行攻击。此时，因为攻击者拥有合法的密钥，加密和认证技术都已经失效，只有通过入侵检测才能发现并清除入侵者。

对无线自组织网络入侵检测方面的研究相对较少，主要集中于提出一些入侵检测的架构，因为无线自组织网络自组织和无中心的特点，集中式的入

侵检测无法应用，设计了一些分布式合作的入侵检测方案。在入侵检测方法上，主要采用异常检测，事先需要进行数据训练。本文的创新点在于，提出一种基于时间自动机的入侵检测方案，将分布式合作的入侵检测架构与时间自动机的检测方法相结合，形成一个适合于无线自组织网络环境的，无需进行数据训练并能实时检测入侵的整体系统。

本文首先指出了无线自组织网络按需路由协议 DSR 的弱点和针对 DSR 的一些攻击方式，然后提出基于时间自动机分布式合作的入侵检测系统。整个入侵检测系统由两部分组成：

（1）分布式合作的入侵检测架构。无线自组织网络具有自组织无管理中心的特点，因此必须采用分布式入侵检测的方法，即入侵检测点分布于整个网络。但为了节省网络资源，又不能所有节点都为入侵检测执行节点，所以，我们提出一种基于簇头的分布式合作的入侵检测，在每一个区域内选出一个簇头节点作为入侵响应的监视节点，负责整个区域节点行为的监视，同时各个监视节点又相互合作检测整个网络节点。所有监视节点形成了对整个网络的入侵检测。

（2）基于时间自动机的入侵检测算法，我们将按需路由协议 DSR 的规范形成时间自动机，节点的行为使用时间自动机进行分析，如果不符合时间自动机的行为则认为是攻击行为。该检测算法不需要事先知道入侵行为的特征，也不需要事先进行数据训练，就可直接进行检测。

本文其余部分如下安排，5.2 节介绍了无线自组织网络入侵检测方向的研究进展。5.3 节介绍一些背景知识，包括 DSR 概述、DSR 的弱点及针对 DSR 的攻击等。5.4 节描述我们提出的入侵检测系统。5.5 节进行了模拟实验，评估了入侵检测系统的效率。5.6 节是结论。

5.2 相关工作

Zhang Yongguang 和 Lee Weeke 提出了一个基于代理的分布式协作入侵检测方案[5]。在该方案中 IDS 代理运行于网络中每一个节点上，拥有六大功能模块，分为数据收集、本地检测、合作检测、本地入侵响应、全局入侵响应、安全通信。图 5-1 为 IDS 代理由六大功能模块组成的示意图。其过程为首先执行本地数据收集和检测。如果本地节点能够确定入侵已

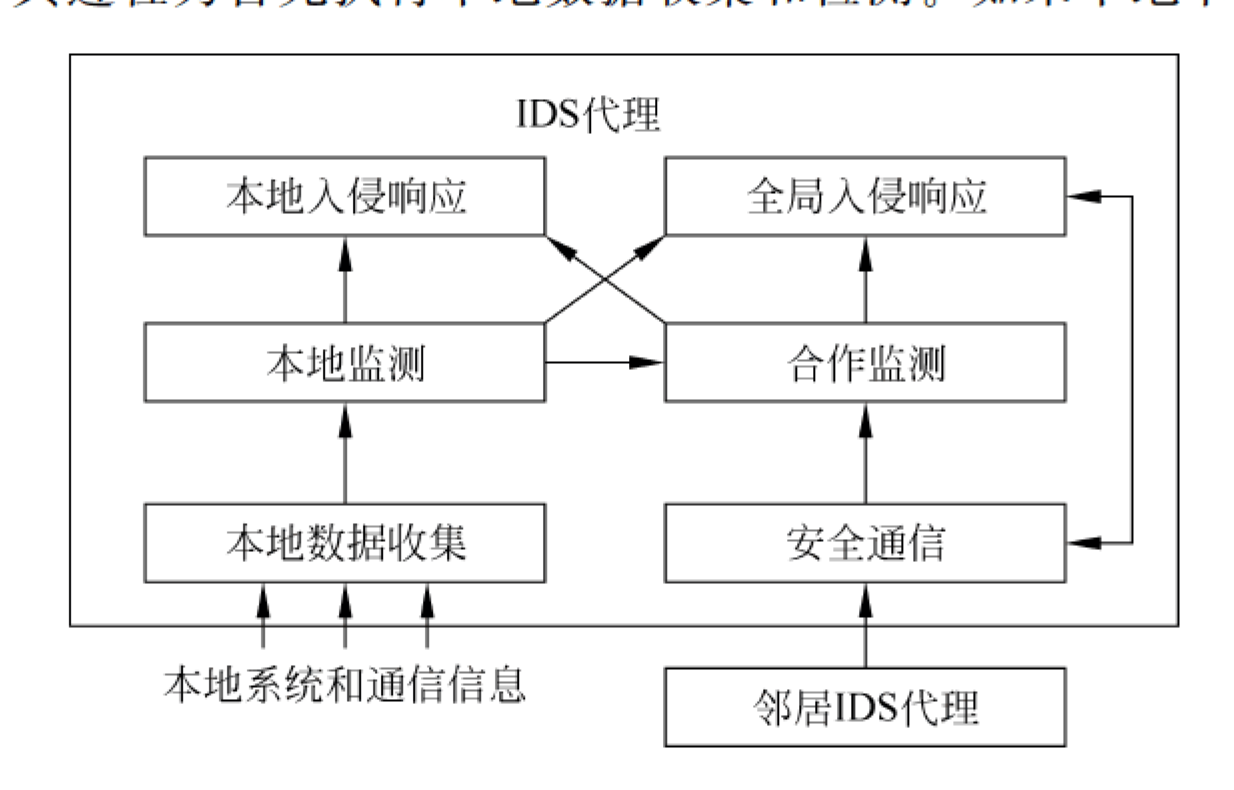


图 5-1　IDS 代理组成

发生，则直接告警。如果只是怀疑有入侵行为，本地节点能够激发多节点的协作检测，进一步检测是否发生了入侵。如果确定有入侵则激发全网的入侵响应。同时提出了一个检测路由进攻的异常检测模型，通过提取正常网络运行时的数据，进行分类训练，实现对路由入侵的检测。为了提高检测效率，入侵检测并不局限于网络层，而是多层综合检测。Zhang Yongguang 在文献[6]中对上述方案进行了详细的论述，建立了一个 IDS 模型并用网络模拟器实现了模拟运行。

上述方案的优势主要有两点：其一，提出了分布式协作入侵检测的架构，利用分布在每个节点的 IDS 代理独立完成本地检测，合作完成全局检测，适合于无线自组织网络自组织的特点；其二，采用多层综合入侵检测，提高了检测效率。缺点也主要有两点：其一，采用异常检测模式，要事先采样数据进行训练，不适合于无线自组织网络多变的应用场合；其二，每个节点都运行有代理，占用过多的内存和计算资源。

Oleg Kachirski 和 Ratan Guha 提出了基于移动代理的入侵检测方案[7]。他们认为 Zhang Yongguan 的方案每个节点都有代理，过于占用网络资源，为了节省资源，只是在某些节点上驻留有监视网络的代理，并且代理的数量可按要求进行增减。

Tseng Chin-Yang 等人提出了基于规范(specification-based)入侵检测方案[8]。该方案利用分布在网络中的监测点，合作监视在 AODV 路由查询过程中，被监视节点是否按路由规范进行操作。如果发现不一致则报警。检测过程为，监听节点对查询报文的处理过程，记录下来形成转发表和操作树，然后用规范形成的有限状态机进行检查，输出为正常状态、怀疑状态、入侵状态三种结果，再分别进行不同的处理。该方案优点在于采用了基于规范入侵检测，因此可以既不需要事先提取入侵行为特征，也不需要数据进行训练，有较高的检测率和较低的误报率。主要缺点为，占用节点较多的计算资源，也未用实验进行验证。

R. S. Puttini 等人设计了一种分布式的入侵检测架构[9]，该架构使用基于特征的入侵检测技术。Huang Yi-an 和 Lee Wenke 提出合作检测的系统[10]，该系统通过一些简单的规则来识别入侵者。B. Sun、K. Wu 和 U. W. Pooch 设计一种入侵检测代理[11]，该代理利用马尔可夫链来进行入侵行为识别。P. Albers 提出一种利用简单网络管理协议(SNMP)所使用的管理信息库(MIB)作为入侵检测源数据的架构[12]。S. Bhargava 和 D. P. Agrawal 提出一种入侵检测和响应的模型[13]。Weichao Wang 提出一种鉴别 AODV 协议中序列号伪造的方法[14]。表 5-1 对三种主要的入侵检测方案进行了比较。

表 5-1　三种入侵检测方案的比较

协议名称	基于代理的分布式协作入侵检测	基于移动代理的入侵检测	基于规范入侵检测
执行者	驻留节点上的代理	各种移动代理	每个节点
检测模式	异常检测	异常检测	基于规范的检测
检测方法	分布式监测、邻居监视	分布式监测、邻居监视	分布式监测、邻居监视
优点	各代理合作监测与响应	可动态调整代理数量，降低对资源的消耗	不需要数据进行训练，较高的检测率
缺点	占用过多资源	协议比较复杂	计算量大

通过上述比较可以看出，上述入侵检测方案存在以下一些特点：

(1) 现行的入侵检测的架构为使用代理作为入侵检测的执行者，代理驻留并运行于网络中每一个节点内，分布式的监视网络状况，信息共享，合作检测入侵行为。这种架构对于入侵检测来说是较为有效的，但未充分考虑到网络带宽和节点计算资源有限的特点，节点本身要运行自己的应用程序，又要负责网络报文转发，CPU 和内存资源已经很紧张，还要运行代理监视网络和主机，这也许会导致节点性能明显下降，甚至资源枯竭。网络带宽也十分有限，代理间信息交换也要占用带宽，这也许会影响网络性能。我们认为在设计上应充分考虑到网络资源有限的特点，降低其对资源的要求，不必每一个节点都运行代理，可采用两种方式，一种是分区域，每个区域使用一个代理负责监控。另一种是使用少量移动代理散布于网络各处，如发现异常，可向异常处移动，进一步检测以确定是网络故障还是入侵行为。

(2) 入侵检测的模式通常为异常检测模式。异常检测模式是定义正常行为的范围，凡是偏离了正常的行为都为入侵行为。该检测模式因为其能够检测出新的入侵行为而被采用。但在无线自组织网络环境下，因为动态的拓扑、无线信道的不稳定、应用环境的多变，使得难以准确定义正常行为的范围。如果定义不当，将会导致不能发现入侵行为或者错误报警率太高。针对此不足，应该采用基于规范的入侵检测模式，事先按网络协议的规范定义好程序的执行步骤，运行时监视程序是否按规范执行，偏离了正常行为即为入侵行为。无线自组织网络协议都有明确的规范，使用该技术既能检测出未知的入侵行为，又有较低的误报警率。

5.3 背景知识

5.3.1 DSR 概述

DSR(Dynamic Source Routing)路由协议[15]是较为经典的无线自组织网络路由协议，它是一种按需路由协议。若源节点 S 需要给目标节点 D 发送数据报文，但它的路由缓存中并没有从源 S 到达目标 D 的路由，此时节点 S 先将数据存入它的数据缓存区，再以广播方式向周围节点发送路由查询报文(route request)，每个路由查询报文通过序列号和源节点来唯一标识。周围节点收到路由查询报文后，如果它以前收到并处理过同样的报文，则直接抛弃不处理。如果没有收到过该报文，节点把自己的地址添加到路由发现报文的地址列表，并向周围广播转发。当路由查询报文到达目标节点或中间节点具有到目标节点的路由，该节点把路由发现记录的地址信息再加上自己的地址信息结合生成路由回答报文(route reply)，单播发送回源节点 S。当源节点收到路由回答报文时，存储该路由信息，用于数据报文的发送。

路由维护负责监测网络中正在使用路由的通断情况，并随时通知源送节点有关该路由出现的错误情况。当网络中的某一节点按照报文中的路由信息进行报文转发，出现无法将报文继续转发到下一节点的情况，并且经过多次重发无效后，它就产生一个路由出错报文(route error)来通知源节点目前使用的路由已经失效，路由出错报文指出了出错的链路，即产生路由出错报文的节点地址和不能到达的下一跳的节点地址。源节点和其他的节点一旦收到路由出错报文，就会检查自己路由缓存中的路由并且删除出错的路由，同时，源节点会

立刻查询自己的路由缓存以寻找一条替代路由使用，若找不到替代路由，则重新激发路由请求报文寻找目的节点。

5.3.2　时间自动机简介

在计算机科学领域内，自动机是分析模拟许多现象的有利工具，特别是在并行有限状态系统中，自动机理论有着很重要的地位。在并行计算的跟踪模型计算中，系统是由其行为来区别的。设由一系列状态或事件来代表行为，系统的行为集合就是形式语言。由此，系统就可用产生该语言的自动机来模拟。从而复合或复杂系统就可通过产生模拟其子系统的自动机来模拟。为捕捉实时系统的行为，计算模型需要用时间概念扩展。为此，时间自动机提供了一套简单有效的方法。使用有限个变量来表示时间，称为时间变量。同时用一个条件来注释状态转换图。由于这个与时间有关的条件决定了状态的转换发生的时机，因此称之为时间限制。对于某个字符串，每个字符对应一个实型的时间标志称之为时间字。对某次转换，时间自动机可能检查其时钟值，并对一些时钟赋以新值。那么时间自动机就可以接受时间字。自动机理论可以对实时系统用有限控制解决一些验证问题。本文就采用时间自动机对节点的行为实时地按路由协议规范进行验证。

5.3.3　DSR的弱点和攻击方式

动态源路由协议 DSR 中没有考虑任何安全的防护措施，其假定参与路由协议信息交换的所有节点均能够诚实地转发和处理路由信息，这导致 DSR 容易遭受各种安全攻击。在 DSR 路由协议中，一些关键的数据字段，如源节点地址、目标地址、地址列表，是非常重要的，对其进行任何非法修改都将导致路由的不正常。一个入侵者可采用下列攻击方式：

- 入侵者假冒另外一个节点的地址发出路由查询报文。
- 当入侵者转发报文时，对报文中的地址列表进行插入、删除或修改。
- 入侵者假冒目标节点编造路由回答报文，发回源节点。
- 全部或部分抛弃路由查询报文、路由回答报文和数据报文。
- 编造路由出错报文，谎称正常的路由已经中断。

上述攻击将会导致下列后果：

- 黑洞：节点不转发任何报文。
- 环路：路由首尾相连形成一个环路，进入环路的报文一直在绕圈子，永远不能达到目标节点。
- 网络分割：物理上相连的整个网络，逻辑上被分割为互相不相连的几个子网，导致许多节点之间不能通信。
- 拒绝服务：节点因为资源被大量占用，不能接收和转发报文。

下面具体介绍几种攻击形式。

1. 修改攻击

所谓修改攻击，就是入侵者恶意地修改、插入、删除经过其转发的报文，导致路由异常。因为源路由协议 DSR 中节点对路由报文中的字段没有认证和鉴别能力，所以对路由报文中

的修改其他节点也无法发现。例如,图 5-2 中有五个节点 A、B、C、D、E,节点 A 是源节点,节点 E 是目标节点,节点 A 要建立一条从 A 到 E 的路由。节点 A 发出路由查询报文,经过节点 B、C、D 转发,到达目标节点 E。路由查找过程中每个中间节点转发时,都将自己的节点加入地址列表,如图 5-2 中上一行箭头所示,箭头上字母表示地址列表。节点 C 收到从节点 B 发来的路由查询报文后,它将自己的节点地址 C 加入到地址列表中,形成新的地址列表 ABC,然后转发,以此类推,到达节点 D 时,又增加一个节点 D,形成地址列表 $ABCD$。到达目标节点 E 时,节点 E 生成路由回答报文。路由回答报文包含路由查询时形成的地址列表 $ABCDE$,并沿原路返回。中间节点收到路由回答时,不对地址列表修改,直接转发。图 5-1 中,第二行箭头所示,其箭头下的字母表示路由回答中的地址列表,在转发过程中不能改变。

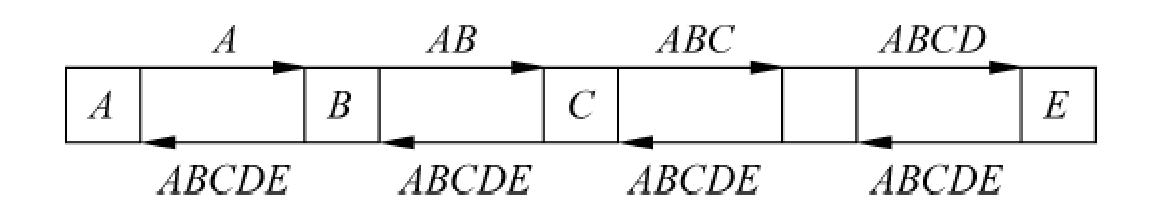


图 5-2　在路由查询和路由回答报文中地址列表的变化

图 5-3 显示一个修改攻击例子,节点 C 是攻击者,当节点收到路由查询时,它按正常协议规范处理,将本节点地址加入路由查询报文中的地址列表中,并转发。路由查询报文到达目标节点 E 后,节点 E 生成路由回答,由来路返回至源节点。当路由查询报文到达节点 C 时,节点 C 将报文中的地址列表 $ABCDE$,删去了一个节点地址 D,形成 $ABCE$ 后转发。节点 B 收到节点 C 转发的路由回答,再转发给节点 A。节点 A 建立一条从 A 到 E 的路由 $ABCE$。但现在此路由已无法正常通信。攻击者 C 还能够删除或修改其他通过节点 C 的报文,导致报文无法正常送达。

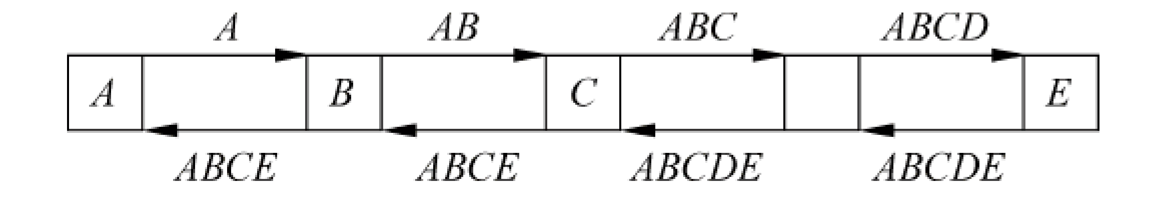


图 5-3　在修改攻击时路由回答报文中地址列表的变化

2. 抛弃报文

在无线自组织网络中的攻击者能够暗中抛弃某些或者全部通过其转发的报文。这种攻击方式容易实施,却难于检测,特别是处于动态的拓扑变化之中时,难于区分是节点移动导致的链路中断、报文丢失,还是故意抛弃造成的报文丢失。

3. 假冒攻击

攻击者假冒其他节点标识或地址发送报文,称为假冒攻击。其方法是将发送报文的源节点地址填为被假冒的地址,而不是攻击者的地址。例如,图 5-2 中,节点 C 可假冒节点 A 发送大量攻击报文到节点 E,节点 E 不能判别报文的真实来源,认为是节点 A 发送的攻击报文。结果,攻击者 C 既攻击了节点 E,又能够隐藏自己的身份、嫁祸于人。

4. 编造攻击

攻击者在网络中编造并散布假的路由信息,称为编造攻击。在动态源路由协议 DSR 中,当一条正在通信的路由中断时,其中断链路的上游节点应该产生一个路由出错报文,返

回到源节点。这份路由出错报文,通知沿途各节点和源节点,将这条已经中断路由从节点路由表中删除。攻击者可以利用这种机制发动 DoS 攻击。例如,在图 5-2 中,从源节点 A 到目标节点 E 有条路由 $ABCDE$ 是正常的。此时,攻击者 C 编造路由出错报文,宣称 CD 之间路由已经中断,并将路由出错报文发给上游节点 B。节点 B 收到此报文后,认为 CD 之间路由中断,将其路由表中路由 $ABCDE$ 删除,并将路由出错报文继续转发到节点 A。节点 A 收到后,也将路由 $ABCDE$ 删除。这条本来正常存在的路由就不能通信了。

5.4　入侵检测算法

5.4.1　监视节点选举算法

在无线自组织网络中,节点的资源是有限的,如果将每个节点都作为入侵检测节点,是非常耗费网络节点资源的。为了节省节点资源,我们提出基于簇头的监视模式,整个网络划分为一个个区域,每个区域选出一个簇头作为监视节点负责整个区域入侵检测。该簇头收集整个区域内的节点的行为信息,并按路由规范进行分析,确定入侵行为。

选举算法由两部分组成,选举阶段和维持阶段。在选举阶段,随机而竞争性地选出监视节点。起始时,整个网络没有一个监视节点,在一段时间内如果没有任何监视节点的信息,任意一节点可以广播一份告示报文宣称自己是监视节点,任何收到此告示报文节点就成为被监视节点,不能再发告示报文。告示报文只能在一跳范围内传播,不能被转发。因为通信是双向的,某个节点能收到告示报文,那么它所发出的报文也能被监视节点收到,所以监视节点能够监视告示报文传播范围内的节点行为。当区域内选举出一个监视节点后,就进入了维持阶段,监视节点周期性地广播告示报文,以维持其监视节点的地位。监视节点服务时间到了后,就重新启动一个新的选举过程,为了保证公平和随机性,上一届的监视节点将不能参加下一届监视节点的选举,除非整个区域只有它一个节点存在。图 5-4 显示三个监视节点及其监视区域分布。

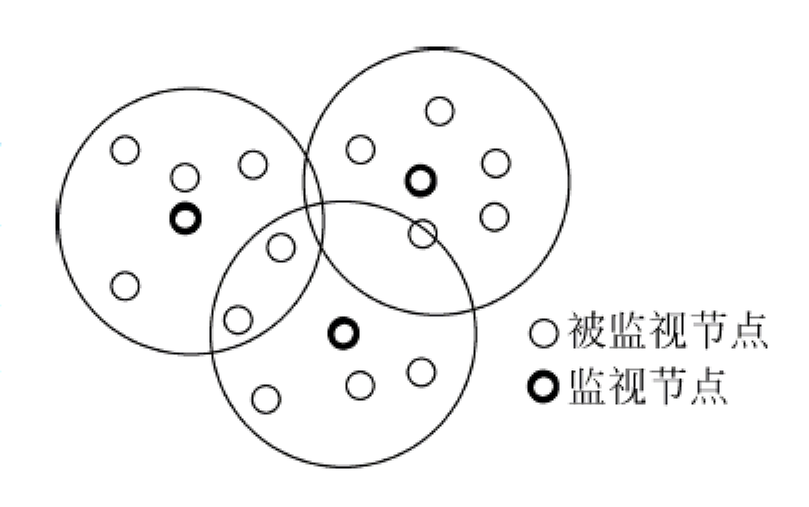


图 5-4　监视节点及监视区域

监视节点的选举是公平而又随机的。所谓公平性,即每个节点都能够有公平的机会选为监视节点,同时每个节点有相同的服务时间。公平性意味着选举的随机性。每个监视节点相同的服务时间要求周期性地重新选举新的监视节点。随机地选举和周期性地更换监视节点,保证了检测的安全性。如果有某个节点是入侵者,又被选举为监视节点,那么在其作为监视节点的期间可以攻击网络而不被发现,因为它是这个区域内的唯一入侵检测点。但它的监视服务时间结束后,又会选出新的监视节点,此时就会发现入侵者。

无线自组织网络中节点可任意移动,监视节点和被监视节点都可能移动而离开原来的区域,如果一个节点在一时间内收不到告示报文,它就可以启动一个选举过程,发出告示报文,宣称自己是监视节点。如果两个监视节点靠近相互收到了告示报文,就比较它们的 ID,那个 ID 较小的节点继续保持为监视节点,另外一个就转变为被监视节点。

5.4.2 基于时间自动机的检测

每个监视节点使用时间自动机来分析监视区域内被监视节点的行为是否符合路由规范并与其他监视节点交换监视信息，以保证节点移动过程中监视的连续性。在监视过程中，它对所监视的每个节点建立时间自动机进行分析。在 DSR 路由协议中，节点可能收到并处理四种类型的报文：路由查询报文、路由回答报文、路由出错报文和数据报文。我们首先对路由查询报文按 DSR 路由规范形成时间自动机如图 5-5 所示。

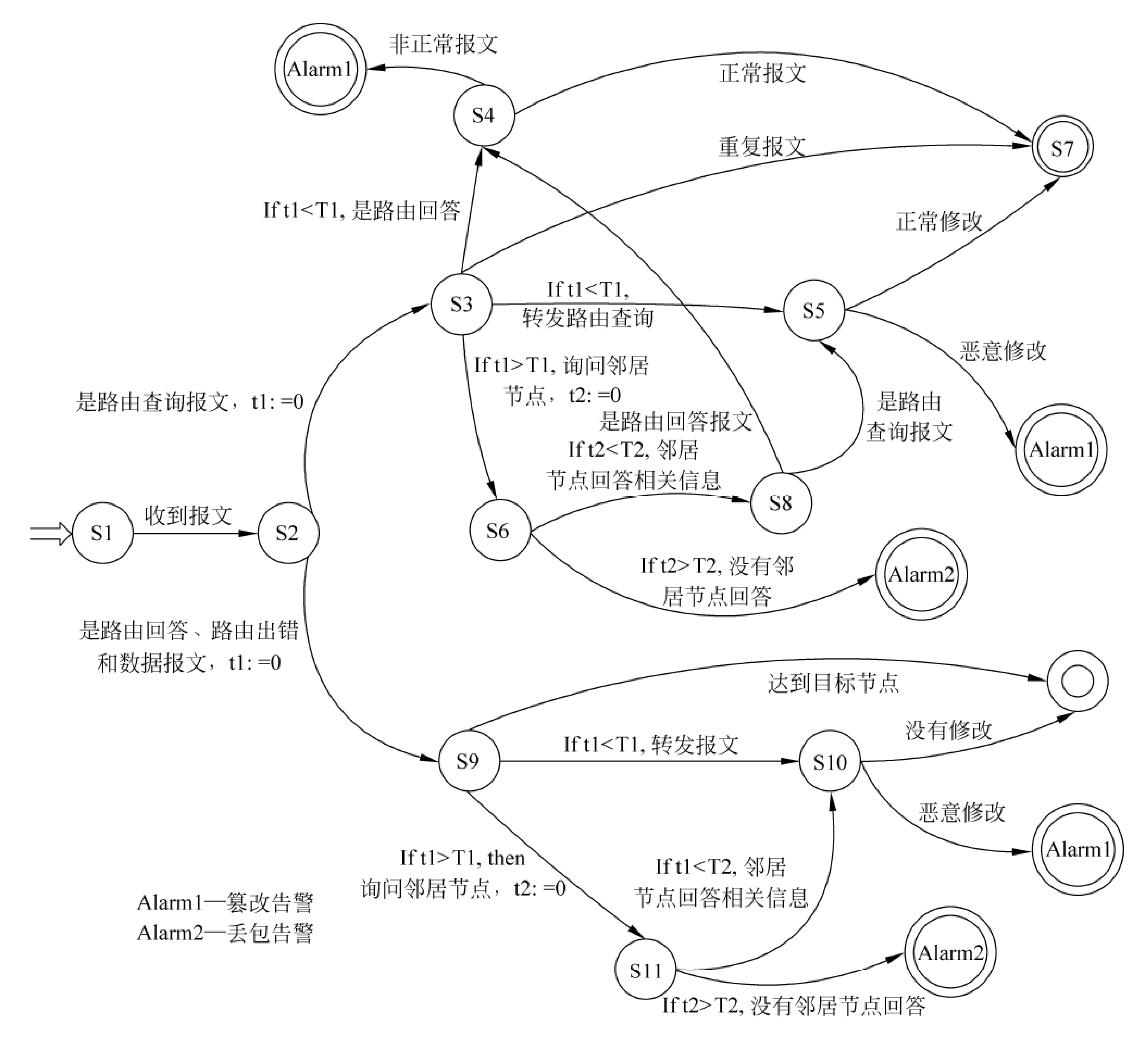


图 5-5　节点收到报文后处理过程的时间自动机

起始状态是 S1，当节点收到报文后，时间自动机转向状态 S2。接下来，时间自动机对接收的报文进行判断，分为两类进行处理，一类是路由查询报文，另一类是路由回答、路由出错和数据报。如果收到的报文是路由回答、路由出错和数据报，进入状态 S9，同时将时钟 t1 设置为 0。如果收到的报文是路由查询报文，进入状态 S3，同时将时钟 t1 设置为 0。这里设置一个有限的时间长度 T1，当节点在 T1 时间内不回答或转发该报文时，就认为该节点抛弃了报文。如果该节点在 T1 时间内产生了路由回答报文，则进入状态 S4，接下来对路由回

答报文按 DSR 路由规范进行检查，如果是正常的，则达到状态 S7，时间自动机正常结束。如果报文有些字段被非法修改了，则发出非法修改告警。在状态 S3 时，如果收到的是以前收到过的路由查询报文，则直接抛弃报文进入终止状态 S7。如果节点在 T1 时间内转发路由查询报文则进入状态 S5，接下来对转发的路由查询报文按 DSR 路由规范进行检查，如果修改了一些不能变化的字段，则发出非法修改告警，否则进入终止状态 S7。如果在状态 S3 超过 T1 时间没有动作，这时有可能是节点移动离开了监视区域，监视节点不能收到节点所转发的路由报文，所以向周围监视节点发出询问，是否收到该节点的转发报文，同时将时间 t2 设置为 0。这里设置另外一个时间长度 T2，用于等待邻居节点的回答。如果邻居监视节点收到并在 T2 时间内将报文信息发回，则状态转向 S8，是路由查询报文将报文信息发到监视节点，转到状态 S5 进行比较，是路由回答报文将报文信息发到监视节点，转到状态 S4 进行比较。如果在 T2 时间内未收到邻居节点报文，说明该节点不参与路由转发，则发出抛弃报文告警。

在 DSR 路由协议中，对于路由回答报文、路由出错、数据报文三种报文的处理过程是同样的，可以采用相同的时间自动机进行分析处理。如图 5-5 所示，起始状态是 S1，当节点收到报文转向状态 S2。如果收到的报文是路由回答、路由出错、数据报文三种报文之一，则进入状态 S9，同时将时间 t1 设置为 0。以下可别转入三个状态，如果本节点已经是报文目标节点，则进入终止状态 S12。如果在 T1 期限内转发报文，则进入状态 S10，在 DSR 路由协议中对这三种报文是只能原样转发不能进行修改的，接下来只要对照一下转发前后的报文，如果有不同则发出非法修改告警，如果相同则进入终止状态。在状态 S9 时，如果超过 T1 没有收到转发报文，则向邻居监视节点查询该报文信息，同时将 t2 设为 0。进入状态 S11，如果邻居监视节点在 T2 周期内收到其发出报文，则将信息发来，进入状态 S10，如果邻居没有收到则发出抛弃报文告警。

图 5-5 是处理当一个节点接收报文后的处理流程。图 5-6 显示的是当节点发送报文后的时间自动机处理流程图。当监视节点监视某个节点 *A* 发出的报文时，状态由 S1 到 S2，可能出现下列两种情况之一：

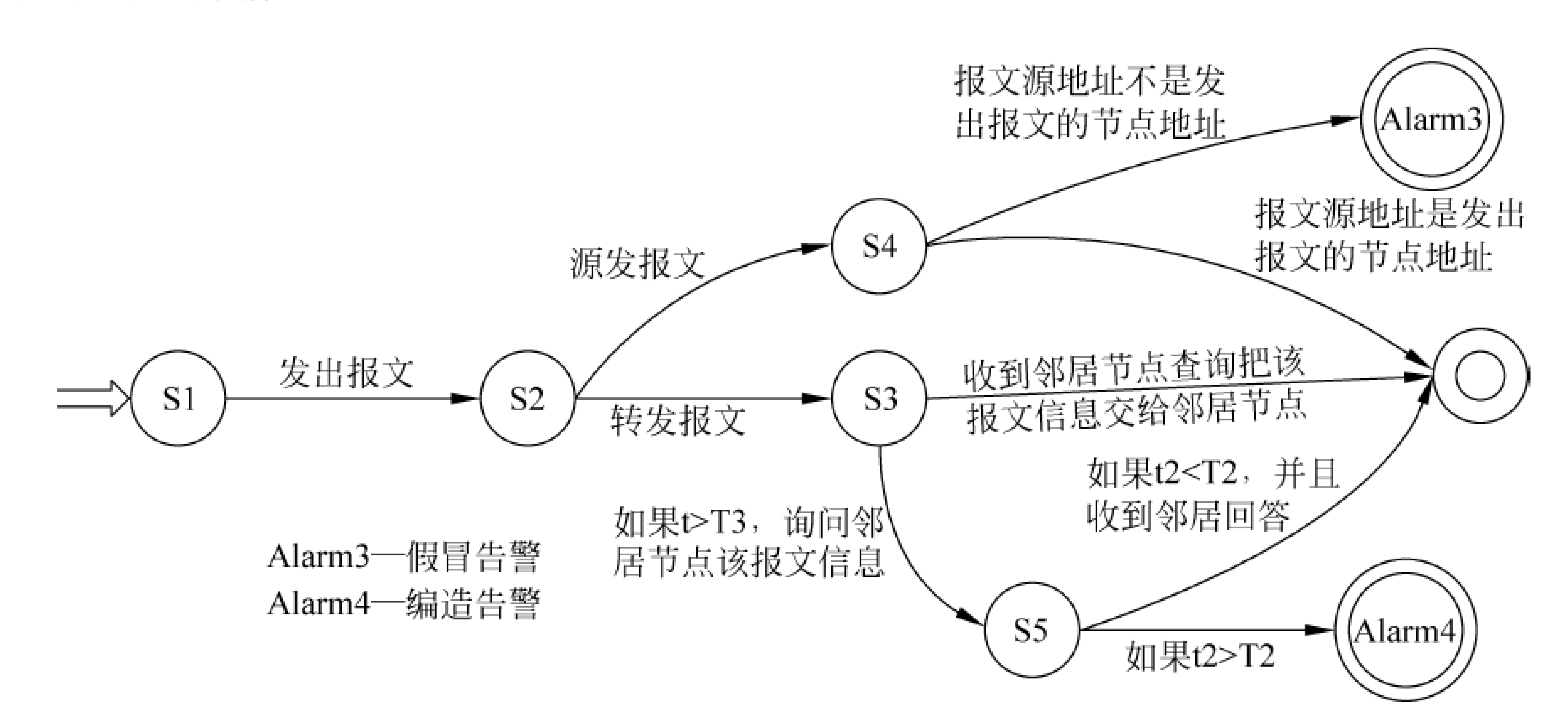


图 5-6　节点发送报文后处理过程的时间自动机

(1) 以 *A* 节点为源节点的报文，即 *A* 节点发出了这个报文，状态是由 S2 到 S4，然后比较一下发出报文源地址与节点的地址，如果相符的话则转入终止状态，如果不相符，则发出

假冒报警。

(2) A 节点是中间节点，它只是转发报文，进入状态 S3，同时设置时间 t3 为 0。可能节点 A 接收报文时，不在监视节点的区域内，转发时节点移动进入监视节点的范围内，所以只看到节点 A 发出了报文，此时监视节点等待邻居监视节点查询，如果在 T3 时间内有邻居监视节点查询，则将报文信息发来邻居监视节点，时间自动机进入终止状态。如果在 T3 时间内没有邻居监视节点查询，那么监视节点主动向周围监视节点发出查询，询问有节点收到过该报文，同时将 t2 设为 0。如果在 T2 时间内没有回答，说明没有节点发送过此报文，是节点 A 编造了此报文，时间自动机发出编造告警。

5.5 模拟实验

5.5.1 实验设置

实验平台为 Pentium 4 配置为 CPU 主频 1.8GHz，RAM 容量 512MB，使用的操作系统是 Red Hat Linux 7.12，仿真平台是 ns-2 2.26(Network Simulator Version 2.26)[31]。仿真中，节点总数设置为 50 个，节点运动范围 1500m×300m，运动速度 0～20m/s，网络中节点的运动采用随机运动模型，即每个节点在该区域内从一点向另一点运动，运动速度在[0,20m/s]内均匀分布，到达目标点后，停留一段时间，然后选择一个新的目标点，同时再选择一个新的速度，向新的目标点运动，以此类推，直至仿真结束。MAC 层使用的 802.11，传输半径为 250m，链路带宽为 2Mb/s。模拟时间为 900s。

5.5.2 实验结果

我们将选举算法和时间自动机在 NS-2 中进行了编码实现。为了检验入侵检测效果，按前述的分析，我们设计了四种对 DSR 路由协议的攻击方式。攻击方式 1 是非法修改攻击，即入侵者转发报文时，非法插入、删除和修改报文中的信息。攻击方式 2 是抛弃攻击，即入侵者只收报文，不转发任何报文。攻击方式 3 是假冒攻击，即入侵者假冒其他节点发送各种报文，如路由查询报文、数据报文等。攻击方式 4 是编造攻击，入侵者编造一些由它转发的报文，但实际上源节点并未发出此报文。

表 5-2 显示时间自动机对四种攻击方式的入侵检测率和误报率。从表 5-2 可以看出，对假冒攻击的检测率最高，主要原因是监视节点只需要将发出报文节点地址与报文中地址列表对照一下，如果没有就是假冒攻击，最为简单。抛弃攻击检测率最低，主要原因是监视节点不能直接做出判断，要到邻居节点去查询才能得出结果。编造攻击也是需要邻居监视节点查询才能判断，检测率也较低。但是总的来说，检测率在 80%以上，说明我们的算法还是十分有效的。

表 5-2　四种攻击方式的入侵检测率

攻击方式	检测率/%	误报率/%	攻击方式	检测率/%	误报率/%
修改攻击	91.3	2.9	假冒攻击	97.4	1.3
抛弃攻击	83.7	5.7	编造攻击	88.5	7.2

5.6　小结

在本章中，我们提出了基于簇头的分布式合作的入侵检测架构，整个网络分成一个个区域，每个区域内的监视节点既负责本地入侵检测又合作检测整个网络节点，通过随机选举簇头作为监视节点，并周期性地重新选举簇头，既节省网络资源又保证了入侵检测系统的安全性。在入侵检测架构的基础上，设计了基于时间自动机的入侵检测算法，通过 DSR 路由协议规范形成节点处理过程的时间自动机，对节点的每个报文的处理过程按时间自动机进行分析，实时地发现入侵行为。最后通过模拟实验检测了我们算法的有效性。

参考文献

[1]　Hu Y C. Perrig A. A Survey of Secure Wireless Ad hoc routing. IEEE Security & Privacy Magazine, May-June 2004,2(3): 28-39.

[2]　Aldar Chan. Distributed Symmetric Key Management for Mobile Ad hoc Networks, IEEE INFOCOM'04, Hong Kong, March 2004.

[3]　Manel Guerrero Zapata. Secure Ad hoc On-Demand Distance Vector Routing. ACM Mobile Computing and Communications Review (MC2R), 2002, 6(3): 106-107.

[4]　Papadimitratos P, Haas Z J. Secure Link State Routing for Mobile Ad hoc Networks, IEEE Workshop on Security and Assurance in Ad hoc Networks, In Conjunction with The 2003 International Symposium on Applications and the Internet, Orlando, FL, January 28, 2003.

[5]　Zhang Yongguang, Lee Wenke. Intrusion Detection in Wireless Ad-hoc Networks, in Proceedings of The Sixth International Conference on Mobile Computing and Networking (MobiCom 2000), Boston, MA, August 2000.

[6]　Zhang Yongguang, Lee Wenke. Intrusion Detection Techniques for Mobile Wireless Networks, Mobile Networks and Applications, 2003.

[7]　Oleg Kachirski, Ratan Guha. Intrusion Detection Using Mobile Agents in Wireless Ad hoc Networks, IEEE Workshop on Knowledge Media Networking (KMN'02).

[8]　Tseng C Y, Poornima Balasubramanyam, Calvin Ko, et al. A Specification-Based Intrusion Detection System For AODV, 2003 ACM Workshop on Security of Ad hoc and Sensor Networks (SASN'03) October 31 2003, George W. Johnson Center at George Mason University, Fairfax, VA, USA.

[9]　Puttini R S, Percher J M, Camp L Mé O, et al. A Modular Architecture for Distributed IDS in MANET, Proceedings of The 2003 International Conference on Computational Science and Its Applications (ICCSA 2003), Springer Verlag, LNCS 2668, San Diego, USA, 2003.

[10]　Huang Yi-an, Lee Wenke. A Cooperative Intrusion Detection System for Ad hoc Networks, 2003 ACM Workshop on Security of Ad hoc and Sensor Networks (SASN '03), Fairfax, VA, October 31, 2003.

[11]　Sun B, Wu K, Pooch U W. Routing Anomaly Detection in Mobile Ad hoc Networks, Proceedings of 12th International Conference on Computer Communications and Networks (ICCCN 03), Dallas, Texas, October 2003: 25-31.

[12]　Albers P, Camp O, Percher J M, et al. Security in Ad hoc Networks: a General Intrusion Detection Architecture Enhancing Trust Based Approaches. In Proceedings of the First International

Workshop on Wireless Information Systems (WIS-2002), Apr. 2002.

[13] Bhargava S, Agrawal D P. Security Enhancements in AODV Protocol for Wireless Ad hoc Networks, Vehicular Technology Conference, 2001, 4: 2143-2147.

[14] Wang Weichao, Lu Y, Bhargava B K. On Vulnerability and Protection of Ad hoc On-demand Distance Vector Protocol, in Proceedings of 10th IEEE International Conference on Telecommunication (ICT), 2003, 16.

[15] David B Johnson, David A Maltz, Hu Yih-Chun, The Dynamic Source Routing Protocol for Mobile Ad hoc Networks (DSR), INTERNET-DRAFT, draft-ietf-manet-dsr-10. txt, 19 July 2004.

第6章　无线自组织网络的主动防护机制

CHAPTER 6

摘要：无线自组织网络是由移动节点自组织形成的网络，由于其动态拓扑、无线信道的特点，安全容易遭受各种威胁。至今提出的许多安全方案主要集中于入侵阻止和入侵检测两个领域内。尽管这些安全方案能够取得一定的安全保障效果，但是它们都只是被动地去发现和阻止入侵者，并不能从根本上消除入侵行为。为了解决这个问题，本文提出一种自动入侵响应模型。该模型通过多种功能的代理（agent）组成一个整体来实现主动入侵响应，首先在每个节点布置监视代理，负责收集其周围每个邻居节点的行为信息。然后每个区域内的决策代理汇总监视代理的信息并进行判断。最后，阻击代理在入侵者周围形成一道移动防火墙，将入侵者包围并隔离于网络，消除入侵行为。并且进行了模拟实验，证实了该模型能够有效地阻止入侵。

关键字：无线自组织网络、安全、入侵检测、入侵响应、移动代理。

6.1　引言

随着计算机网络的不断发展和普及，安全问题日益严重，已成为当今研究的重点。现有防范网络入侵的方法可分为四类，即入侵阻止、入侵检测、容忍入侵和入侵响应。入侵阻止系统利用认证、加密和防火墙技术来保护系统不被入侵者攻击和破坏。但是，实践证明入侵阻止系统并不能防止所有的入侵。入侵检测系统是根据分析采集的主机系统或网络的活动来检测入侵行为，入侵检测系统分为基于主机的入侵检测系统和基于网络的入侵检测系统。有些入侵检测系统只有有限的响应功能，当检测到入侵行为时，它们可以采取报警、通知系统管理员或断开本地连接等多种方式。然而，这些方式大多是被动的，也很难知道入侵者来自何方。人们已经认识到没有任何的系统可以保证绝对不被入侵，所以为了建立一个可生存系统，人们提出了容忍入侵的方法。入侵预防、入侵检测和容忍入侵在解决网络入侵问题上都发挥了很大的作用，但这些方法都是被动地对待入侵问题，不能够主动地反击入侵者，隔绝并阻断其对整个网络的入侵行为。而入侵响应系统在入侵发生后能够主动保护受害系统，阻击入侵者。

目前，入侵响应可以大致分为三类：报警响应、手工响应和主动响应。其中大部分系统采用的是报警响应，报警响应是当检测到入侵行为后，入侵检

测系统向网络系统管理员或相关人员发出告警信息。手工响应是系统提供有限的预先编制好的响应程序，并能指导网络管理员选择合适的程序进行响应。与报警响应相比，这类系统优点明显，但是仍然会给攻击者留下较大的入侵时间窗口。而主动入侵响应系统不需要管理员手工干预，检测到入侵行为后，系统自动进行响应决策，自动执行响应措施。不论是从应对数量惊人的入侵事件考虑，还是从响应时间考虑，主动入侵响应系统都是目前较为理想的响应方法。

Cohen[1]进行一系列网络攻击模拟实验，结果表明：某个精通网络的攻击者试图入侵一个网络，如果发现攻击行为后，10 小时以内没有具体行动去阻止其入侵行为，那么他有 80% 的可能入侵成功。如果响应时间推迟到 20 小时，则入侵成功率将达到 95%。如果响应时间推迟到 30 小时，则入侵者一定会成功。由此可见，能够及时快速的实现响应也是十分重要的，而主动响应就是达到此目标的重要措施。

主动入侵响应系统从响应方式可分为基于主机的入侵响应系统和基于网络的入侵响应系统。基于主机的入侵响应系统主要针对主机系统的入侵，其响应主要在受害主机上进行，其响应方式有记录安全事件、限制用户权限、暂停用户进程、封锁用户账号、建立备份系统等。基于网络的入侵响应系统针对网络的入侵，其响应主要在防火墙、网络设备和网管工作站上进行，其响应方式主要有记录安全事件、隔离入侵者、追踪入侵、断开危险连接、反击攻击者等 IP。主动入侵响应系统从响应范围可分为本地响应系统和协同响应系统。本地响应系统是根据本地的安全事件信息，在本地主机或网络设备上进行局限于本地的响应。协同响应系统是在大规模网络中，各响应系统之间在整个网络内的主机或网络设备上共享安全事件信息、协同响应，使响应系统做出更合理的响应，使系统总的损失达到最小。

与固定有线网络相比，无线自组织网络面临更多的安全威胁。为了保障无线自组织网络的安全，至今已经提出了许多安全解决方案[2]，现阶段的研究主要分为三个方面：

（1）密钥的设置与认证[3~5]。研究在无线自组织网络中无中心自组织的情况下，如何实现密钥的分配，如何实现相互认证，主要采用两种方式，基于门限密钥的管理方案[6]和基于 PGP 的自组织的认证方案[7]。

（2）路由安全方案，研究如何对路由协议提供安全保障，通过对路由协议中的报文提供完整性校验、身份认证等安全措施，防止恶意篡改的发生[8~11]。

（3）入侵检测，研究如何在网络运行中及时发现恶意节点的入侵，通常采用邻居监视、合作检测的方法[12]。入侵检测可分为基于特征的入侵检测[13]和基于异常的入侵检测[14]。密钥的设置与认证和路由安全方案这两种技术可以称为入侵阻止技术，尽管入侵阻止和入侵检测技术在防止入侵方面发挥了巨大的作用，但它们都是被动的防御措施，它们所能取得的最好效果就是防止正常节点成为入侵行为的牺牲者。它们不能有效地消除入侵根源——入侵者。这些入侵者能够继续存在并危害网络系统。为了能够从根本上消除入侵行为，本文提出了一种基于移动代理的主动入侵响应模型。该模型利用代理监视网络运行，当发现入侵者后，通过移动代理在入侵者周围形成一道移动防火墙，将入侵者包围并隔离于网络，最终消除入侵行为。本章的创新之处在于，提出移动防火墙的概念，并将其运用于主动入侵响应。

6.2　相关研究

主动入侵响应在固定网络方面已经进行多年的研究，提出的一些方案大致可以分为三类：基于协议增强技术、基于移动代理技术、基于主动网络技术。基于协议增强技术的主动响应，主要思想是对相关网络协议进行了修改或增加一些协议，使得网络中的路由器、防火墙等能够支持对入侵者的跟踪与响应。相关论文有 D. Schnackenberg 等人提出的一种协同入侵跟踪与响应架构(CITRA)[15]，该架构将入侵检测、防火墙和路由器组成一个整体来追踪入侵源并在入侵者附近阻止入侵行为。其具备的功能为跨越网络边界追踪入侵者、阻止入侵者继续危害网络、报告入侵行为、协调入侵响应。该架构的核心是入侵跟踪与孤立协议(IDIP)[16]。IDIP 将网络分为多个域，每个域内有一个协调管理者。IDIP 协议中对于一次攻击，首先检测到入侵的节点会向它所有的邻居节点 IDIP 发送一个事件报告，接收到的节点会首先判断自己是否在攻击路径上，如果是，它将会继续发送这个事件报告给其他的邻居节点。所有在攻击路径上的节点在向邻居节点 IDIP 发送事件报告的同时，会把这份报告和它已采取的响应发送给协调管理者，协调管理者综合各节点的信息，协调各节点的响应，从而达到全局最优的响应。CITRA 和 IDIP 通过各个网络之间信息的交换，对路由器、防火墙和主机的重新配置，实现跨多个网络对入侵者的自动追踪，最后将入侵者在当地予以孤立。Dan Sterne 对上述架构和协议进行了实验，结果显示其对分布式 DoS 攻击有着良好的防御效果[17]。

基于移动代理技术的主动入侵响应，其主要思想是利用移动代理来实现网络入侵检测和响应功能。Curtis A. Carver 提出一种基于代理自适应技术入侵响应系统框架(Adaptive Agent-based Intrusion Response System，AAIRS)[18]，在 AAIRS 中，入侵检测系统监视整个网络系统并产生安全事件报告。AAIRS 接口代理把事件表示为统一格式，并依据对以往误报率的统计，赋予当前入侵事件一个可信度值，可信度由专家赋值，并在检测过程中采用自学习方式动态更新，将这个值与事件报告交给主分析代理。主分析代理首先判断当前事件是一个新的攻击，还是一个原有攻击的延续；若是一个新的攻击，则创建一个新的分析代理来进行响应分析；如果是一个已有攻击的延续，则将事件报告和可信度交给与原有攻击相对应的分析代理进行处理。分析代理将会调用响应分类代理对事件进行详细分类，据此产生相应的响应策略；分析代理还会调用策略规范代理以保证响应策略符合道德、法律、系统安全政策以及资源和环境等约束；最后分析代理将决策产生的响应策略传给响应决策代理。响应决策代理将响应策略分解为具体的响应步骤，并调用响应工具库中相应的程序来执行响应。在分析代理对响应策略决策的过程中和响应决策代理对具体响应动作的决策过程中，都引入了自适应技术；系统记录每次响应的结果，统计不同策略和具体响应措施在以往事件中成功响应的概率，并据此优化响应的决策过程。

基于主动网络技术的主动入侵响应，其主要思想是利用主动网络的功能，实现对入侵者的查找，并在入侵者附近实现封堵。主动网络由一组称为主动节点的网络节点构成[19]。每个主动节点可以是路由器或交换机，这些主动节点共同构成了主动网络的执行环境。主动网络中的报文，内嵌可执行代码，当其流经主动节点时，报文中的程序被激活执行，从而改变消息自身的内容或改变主动节点的环境状态。当发现入侵后，及时找到入侵源就可以更好

地做出有效响应，如果不能有效地追踪到入侵源，入侵响应只能被动地限制在受害主机附近。例如，当发现攻击时，越是在入侵源附近采取阻断、隔离等措施，响应的效果就越好。所以，入侵跟踪在网络入侵响应中非常重要。根据主动网络的特点，如果采用基于主动网络的入侵响应，可以很好地实现入侵跟踪和自动响应。X. Wang 提出一种利用水印技术追踪入侵行为，并将入侵者进行孤立的算法[20,21]。该算法利用主动网络技术，报文通过路由器时留下一些痕迹——水印，当发现攻击行为时，可以通过激活这些水印，达到实时跟踪入侵者并在当地予以阻断。

上述各种主动响应算法只适用于有线固定网络的入侵检测与主动响应，对于无线自组织网络，由于其动态拓扑、无线信道的特点，使得上述方案无法适用。

6.3 入侵响应模型

移动代理是为了达到某个特定的目标，在对外部环境的相互作用基础上，通过对环境状态的认识以及和其他代理的协作，自主地推进问题解决的处理单位。移动代理具有以下特点：

(1) 自主性。代理拥有内部自治机制和问题解决机制，能够控制自己的行为和内部状态。无需外界的指令即可根据自己的知识和收集到的信息进行判断和行为。

(2) 协作性。代理之间相互通信合作，共同完成某项任务。

(3) 反应性。代理能够根据网络环境的变化做出适当的调整，具有自适应能力。

(4) 移动性。代理能够自主地通过网络从一台主机移动到另一台主机。移动代理具有特点适用于在无线自组织网络中执行入侵响应任务，用它来实现用户的入侵响应系统。

入侵响应系统模型如图 6-1 所示。首先要对网络节点行为进行监视，发现入侵者，实现入侵检测。然后将入侵者及其入侵行为信息交给响应决策组件，做出是否响应的决策。最后由响应执行机构具体实施响应，在执行响应过程中实时监控执行效果，如果入侵行为停止，则停止响应执行。在系统模型中的入侵检测部分在第 5 章已经进行了设计和实现，本章主要讨论响应决策和响应执行。当入侵检测系统发现入侵行为时，它将具体的情况发给决策代理，由决策代理根据决策知识库做出决策。因为主动响应也会造成网络资源的占用和消耗，所以当入侵行为对网络损害较小时，可不启动主动响应。如果做出响应的决定，决策代理将产生阻击代理。阻击代理担任响应执行的任务。这些阻击代理移动并驻留入侵者周围的节点，形成一道移动防火墙将入侵者包围隔离，同时将其链路断开，阻止其任何报文的发送与接收。

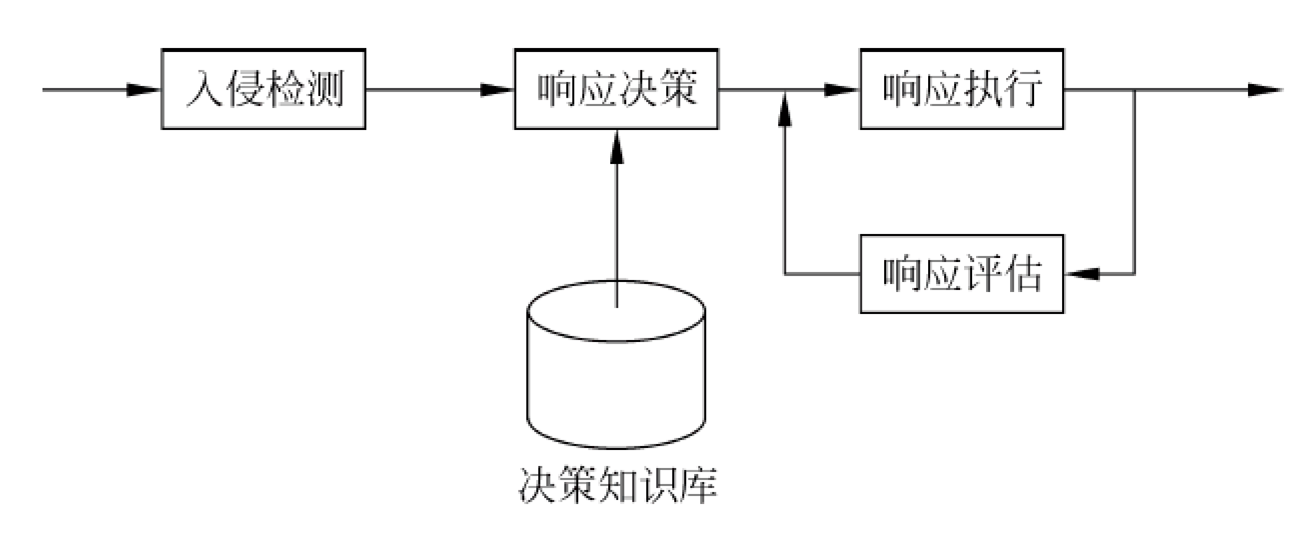


图 6-1　入侵响应系统模型

6.4　主动入侵响应机制

6.4.1　移动防火墙

在固定网络中，防火墙是安装于整个网络的入口处，负责过滤不安全的报文，以保护内部主机免受外部的攻击。无线自组织网络由于节点可随意移动，没有明确的网络边界，所以无法使用防火墙。另外，移动节点可能被截获而成为攻击者，导致攻击从内部产生，防火墙是无法对内部攻击的。因此，我们设计了移动防火墙来阻击内部入侵者。无线自组织网络中节点之间的通信必须借助邻居节点的转发，如果邻居节点拒绝转发报文某个节点，则该节点都就被隔绝于网络。移动防火墙就是利用上述思想进行设计的。当决策代理发现入侵者时，产生阻击代理。阻击代理到达入侵者周围的节点包围入侵者，在其周围形成一道移动防火墙，将入侵者隔绝于网络，阻止其进行攻击。传统的防火墙是用于防止网络之外的攻击，移动防火墙却是用于阻止内部的入侵者。

无线自组织网络中的节点都是移动的，入侵者也可以随意移动。当阻击代理在其周围形成防火墙时，入侵者可以移动而离开包围圈，导致对入侵者隔绝失败。因此，防火墙必须是可以移动的，而且以入侵者的移动为导向，随时将其包围在防火墙之内。为了实现这个目标，用户的移动防火墙由两层组成。第一层是阻击层，由入侵者周围的节点组成，负责包围并隔绝入侵者。第二层是防御层，由阻击层外围的节点组成，负责防止入侵者逃脱。防御层节点平时只是监视入侵者的作用，并不进行阻击，当入侵者移动到防御节点周围时，防御节点就变成了阻击节点，同时在其外围又形成防御节点。为进一步论述如何实现防火墙的移动，下面举一个例子说明，如图6-2所示。

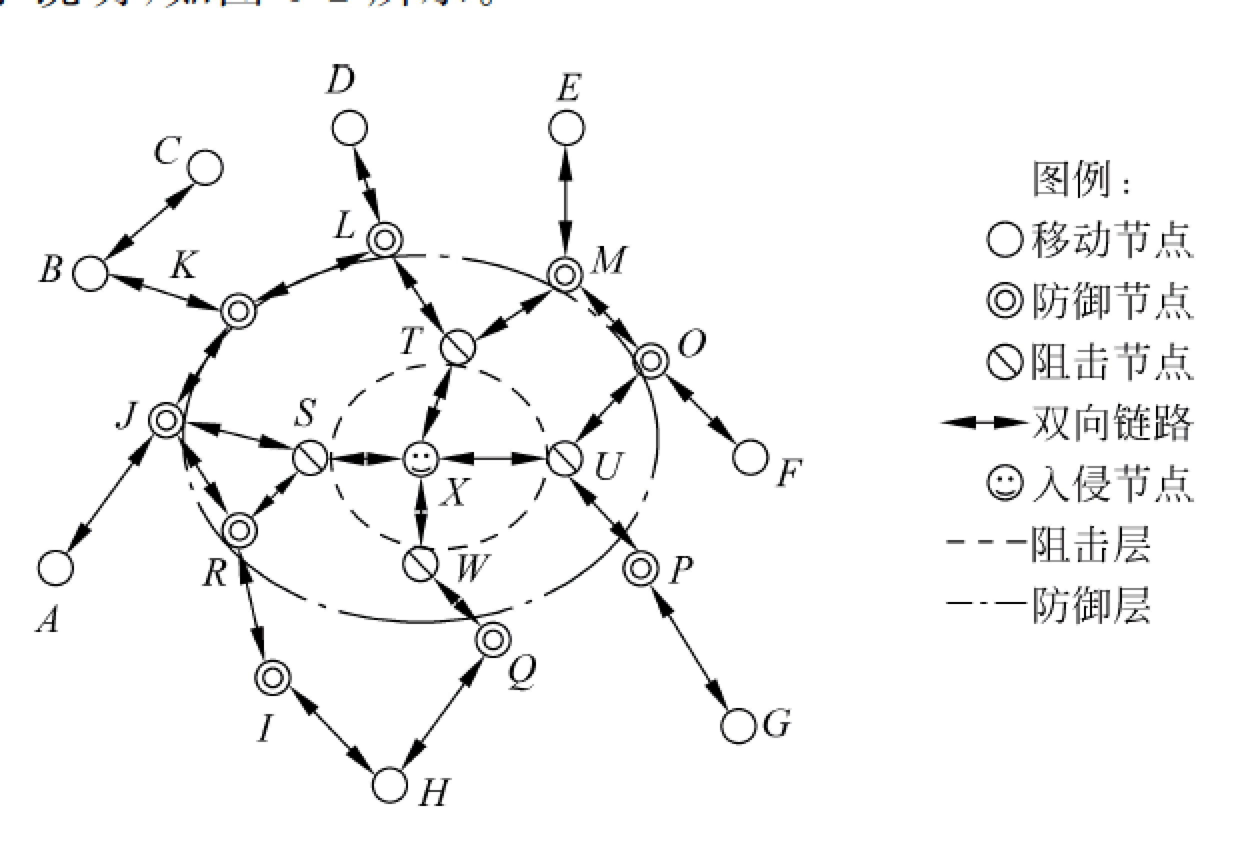


图6-2　移动防火墙和入侵者

图6-2中无线自组织网络由23个节点组成，节点 X 是入侵节点。节点 S、T、U、W 组成防火墙的阻击层，负责切断入侵者的通信链路，隔绝入侵者。阻击层外围节点 J、K、L、M、O、P、Q、I、R 组成防火墙的防御层。当入侵节点移动时，如图6-3所示，节点 K、L 成为入侵节点的邻居，它们就由防御节点转化为阻击节点，其外围节点 B、C、D 又形成新的防御节点。防火墙重新布局，阻击层为节点 K、L、T、S，防御层为节点 J、B、C、D、M、U、W、R，又将

入侵者 X 包围在移动防火墙内部。不同功能的节点,实际是阻击代理在不同节点上发挥不同的作用。当节点与入侵者相邻时,其节点上的阻击代理就发挥切断入侵者的通信链路的作用,该节点就叫阻击节点。当节点是阻击节点的外围时,其节点上的阻击代理就只起监视作用,此节点就叫防御节点。当入侵者移动,防御节点变成阻击节点时,其上的阻击代理会繁殖一些代理移动到外围节点,重新构成防御层。而一些阻击节点由于不再与入侵者相邻会变成防御节点,一些防御节点会变成普通节点,节点 O、P、Q、I 因为入侵者离开而成为普通节点,其上的阻击代理会因超时而死亡。通过上述方法实现了防火墙的实时移动,随时将入侵者包围并隔绝于网络。为了节省资源,移动防火墙并不是永远存在,它与攻击者同存亡。只要入侵者存在并攻击,移动防火墙就存在,但当入侵者资源耗尽,停止攻击时,过一段时间节点上的阻止代理会因攻击源消失而自行消失,移动防火墙也就不存在了。

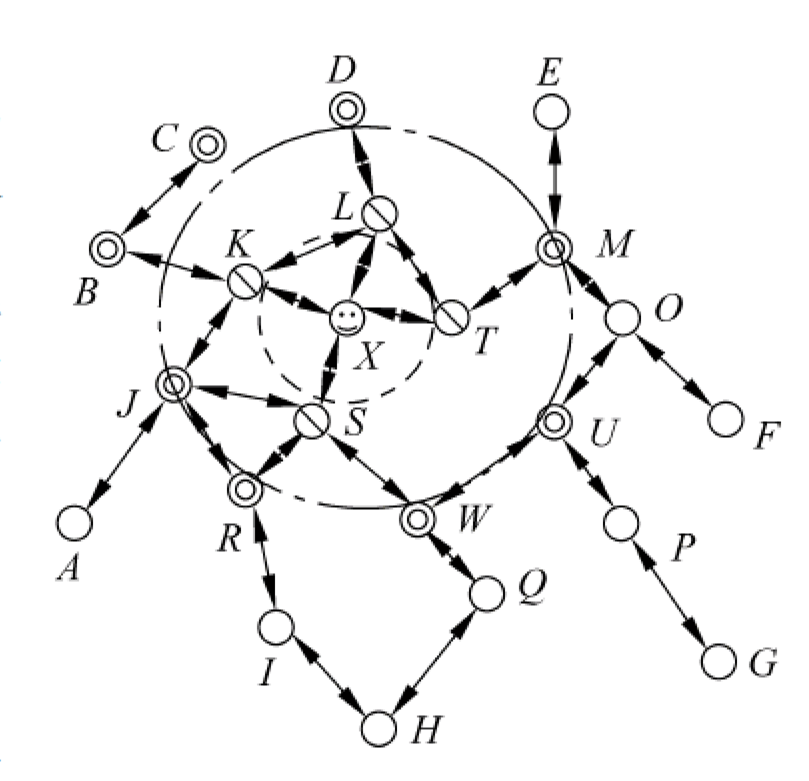


图 6-3　入侵者移动后重新形成防火墙

6.4.2　阻击代理的移动方式

当决策代理发现入侵者时,产生阻击代理将入侵者包围并隔离。但决策代理并不一定就在入侵者周围,可能会有一段距离,此时如果决策代理产生大量阻击代理同时向入侵者方向移动,会过多占用信道,导致网络拥塞。为了减少对资源的占用,我们设计一种节省资源的移动方式,我们称之为开花弹方式,即当炮弹到达目标上空时才爆炸,产生无数子弹来消灭敌人。当决策代理发现入侵者时,也只产生一个阻击代理,向入侵者移动,到达入侵者周围的节点时,这个阻击代理复制许多阻击代理,散开在入侵者周围,形成移动防火墙将入侵者包围并隔离。

6.4.3　本地修复

入侵者节点在没有被发现之前,在网络中参与路径转发,会成为多个路径的中间节点,当其识别并隔离之后,通过入侵者的路径就要被中断。为了减少路径损失,提出本地修复的策略。即当入侵者被隔离后,被中断路径的上游节点,发起一个路由查询,寻找一条绕过入侵者的路径,如果能够找到这样一条路径,将这条路径替换已经中断的路径,就等于将这条路径修复正常了。如果不能发现绕过入侵者的路径,就只能向源节点发送一条路径中断的消息。由源节点进行路由查询了。

6.5　实例分析

为了进一步阐述我们设计的安全架构的功能和特点,本节举一个例子来说明整个入侵检测和响应的过程。图 6-4 为一个无线自组织网络的拓扑图。

图 6-4 中有 15 个移动节点,从节点 A 到节点 O,相邻节点通过双向链路进行连接,其中

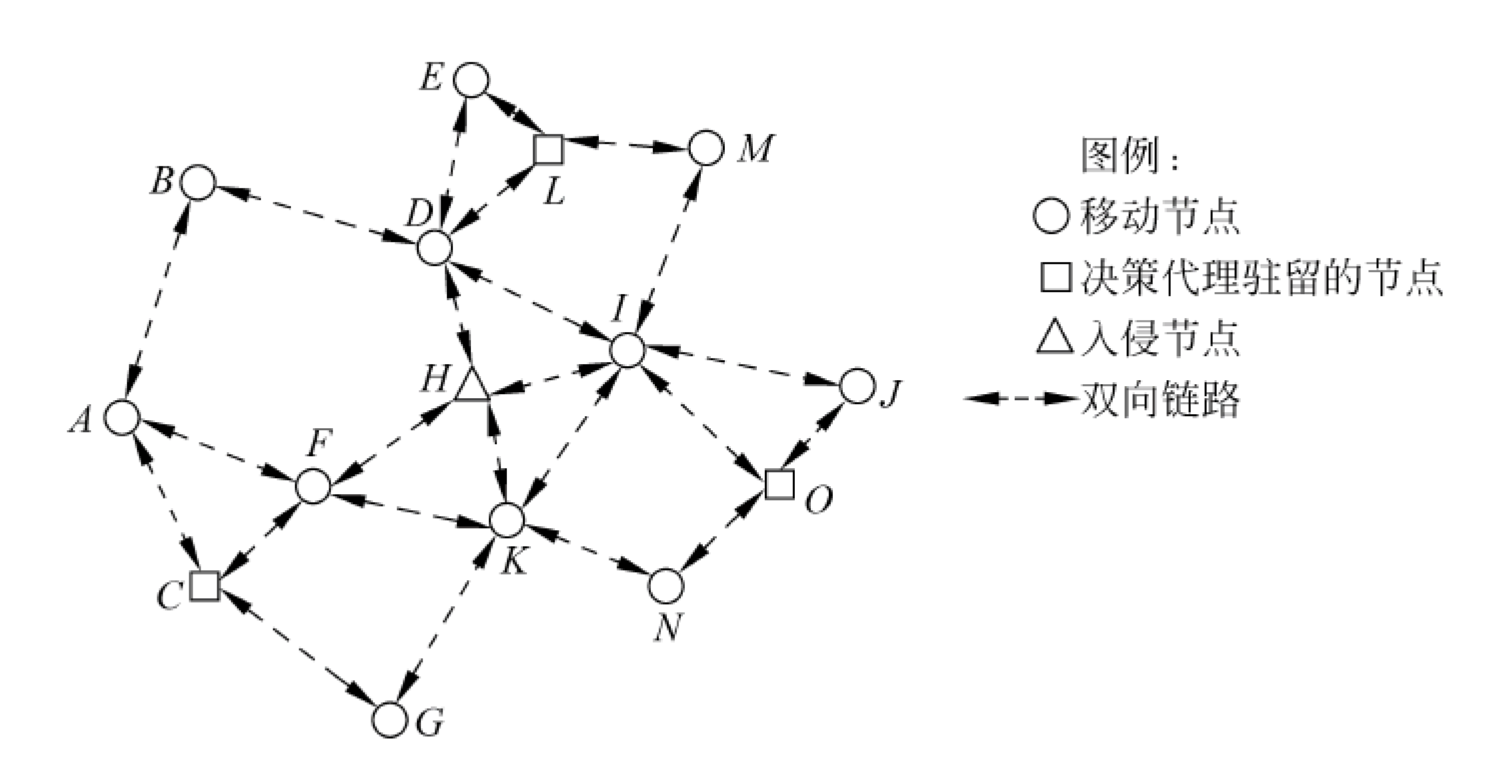


图 6-4　一个无线自组织网络拓扑图

节点 H 为入侵者，每个节点都驻留监视代理，监听并收集其邻居节点的行为信息，三个节点 C、L、O 中驻留了决策代理，负责其区域内信息的汇总与决策，如节点 L 上的决策代理就负责汇总节点 D、E、M、L 节点上监视代理所收集的信息。

图 6-5 显示入侵节点 H 开始发动拒绝服务攻击，它向整个网络泛洪发送大量无用数据报文或路由查询报文，数据报从入侵者周围节点开始向整个网络扩散，大量占用和消耗网络资源，导致其他节点无法正常传送报文。

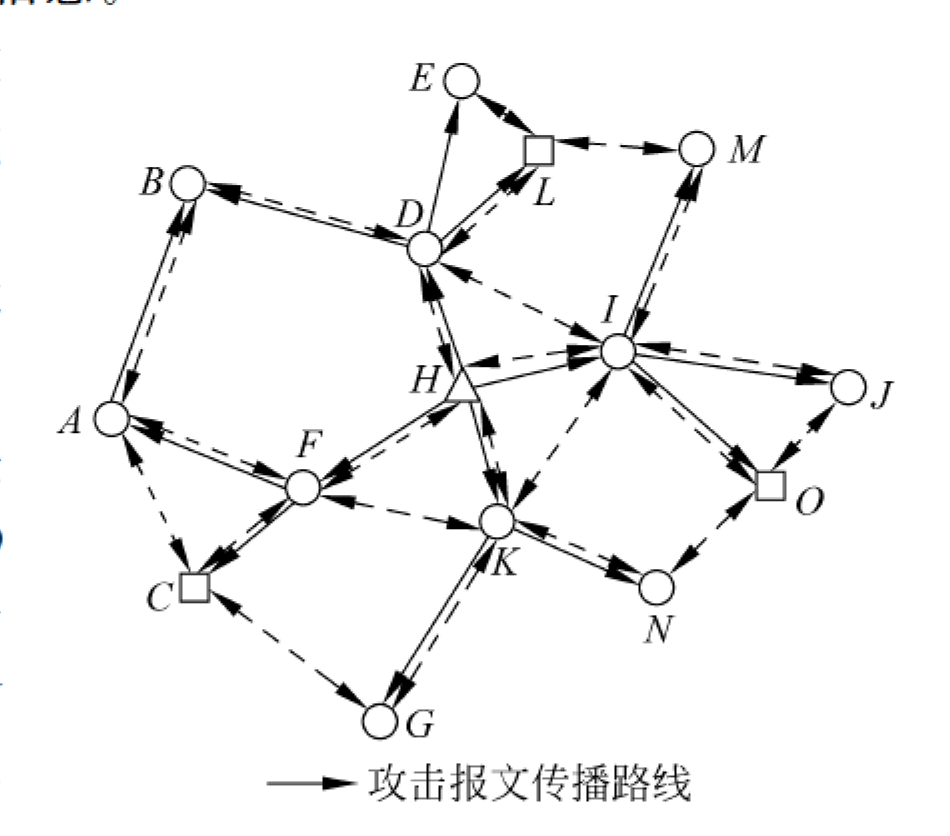


图 6-5　入侵节点发动攻击

整个安全架构的响应可分为入侵检测和入侵响应两个过程，首先是入侵检测，节点 F、G、I、D 是 H 的邻居，在节点 F、G、I、D 上的监视代理时刻监视节点 H 的行为并将其行为进行编码，当 H 节点连续发送查询报文时，F 节点上的代理将编码发往 C 节点上的决策代理，D 节点上代理的监视数据发向 L 节点，G、I 节点上代理的监视数据发往 O 节点。决策代理调用策略库中的路由规范进行判断。判断为入侵行为后，下一步进行入侵响应，决策代理的响应模块开始产生阻击代理。在节点 C、L、O 上的决策代理判断有入侵后，分别产生阻击代理。节点 C 上决策代理产生的阻击代理沿 CF 链路到达入侵者 H 的邻节点 F，到达后将节点 F 与入侵者 H 的链路 FH 中断，拒绝 H 节点的任何路由报文。同样，节点 L 和 O 上的决策代理产生的阻击代理分别到达入侵者的另外三个邻居节点 D、I、G，同时将其与节点 H 的链路 DH、IH、GH 断开。这样入侵者 H 虽然在网络中，但已完全被其周围节点隔离。如图 6-6 所示，阻击代理移动到入侵者周围四个节点驻留，形成一道移动防火墙，如图中的虚线，将入侵者隔离。图 6-6 显示阻击代理的移动和孤立的过程。

从上述分析可以看出，遍布整个网络的监视代理实现对每个节点的监控，将节点的行为编码后发送到决策代理，决策代理进行判断。如果发现入侵者，则决策代理产生阻击代理，由阻击代理将入侵者包围并隔离，最终消除入侵的影响，实现网络的正常运转，整个过程是自动进行的，无需人工干预，实现了实时的主动入侵响应。

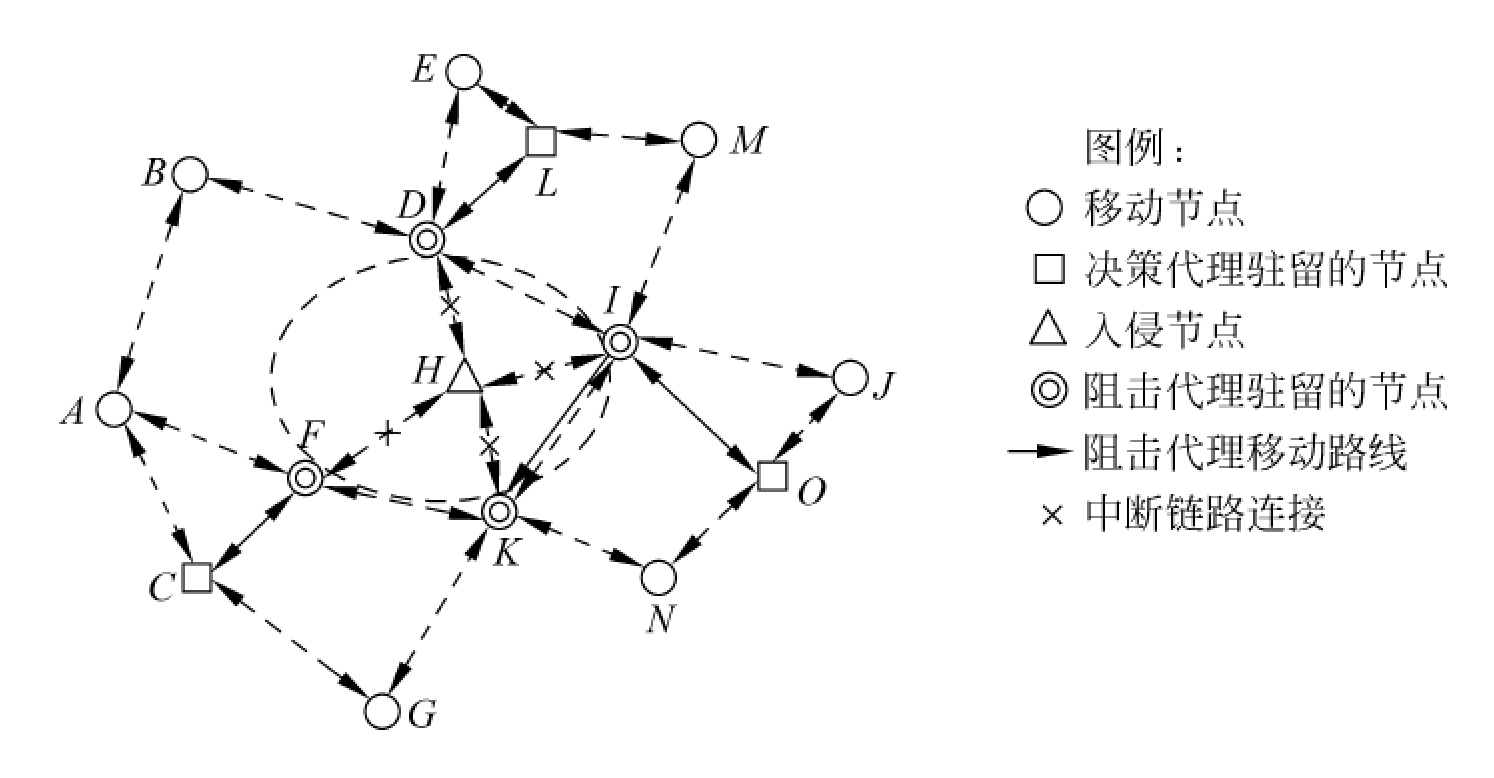


图 6-6　阻击代理包围并孤立入侵节点

6.6　模拟实验

6.6.1　实验设置

实验平台为 Pentium 4，配置为 CPU 主频 1.8GHz，RAM 容量为 512MB，使用的操作系统是 Red Hat Linux 7.2，网络仿真平台是 ns-2 2.26（Network Simulator Version 2.26）。仿真中，节点总数设置为 50 个，节点运动范围 1500m×300m，运动速度 0～20m/s，网络中节点的运动方式采用随机运动模型，即每个节点在该区域内从一点向另一点运动，运动速度在[0，20m/s]内均匀分布，到达目标点后，停留一段时间，然后随机选择一个新的目标点和一个新的速度，向新的目标点运动，以此类推，直至仿真结束。MAC 层使用的 802.11，传输半径为 250m，链路带宽为 2Mb/s。模拟时间为 900s。

测试收集以下两种数据：

(1) 分组传递率(packet delivery rate)。应用层信源发送的分组数目与信宿接收分组数目之比。它描述的是通过应用层观察到的报文丢失率，又反映了网络所支持的最大吞吐量。它是路由协议完成性和正确性的指标。

(2) 平均延迟(average delay)。它是报文从源节点到目标节点的平均传输时间，它反映了网络性能。

6.6.2　实验结果

在实验中，监视代理驻留在每个节点。决策代理不是在每个节点上都驻留，而是划分区域，每个区域内有一个决策代理。阻击代理在节点之间的移动实现起来较为困难，在模拟实验中进行了简化，阻击代理首先驻留在每个节点上，处于休眠状态，当需要阻击代理移动时，只要决策代理发一个报文到节点上，就可以激活阻击代理。实验前 200s 时，没有入侵者。从第 200s 开始，入侵者开始攻击，其后主动入侵响应开始，阻击代理形成移动防火墙将入侵者包围并隔离。每 100s 统计一次分组传递率和传输延迟。从图 6-7 和图 6-8 实验结果可以看出，受到攻击时，分组传递率明显下降，最低时只有大约一半的报文能够传输到目标节点。主动入侵响应开始后，入侵者被包围并隔离，分组传递率开始上升，最终达到 90%左右。虽

然防火墙隔离了入侵节点，消除了攻击行为，但防火墙本身及其移动过程也会占用节点和通信资源，所以分组传递率无法达到没有入侵者时的水平。从图6-8也可以看出，有攻击时，报文传输延迟大大增加，入侵响应开始后，传输延迟缓慢下降，最终达到200ms左右。从实验结果证实移动防火墙能够有效地阻止入侵行为。

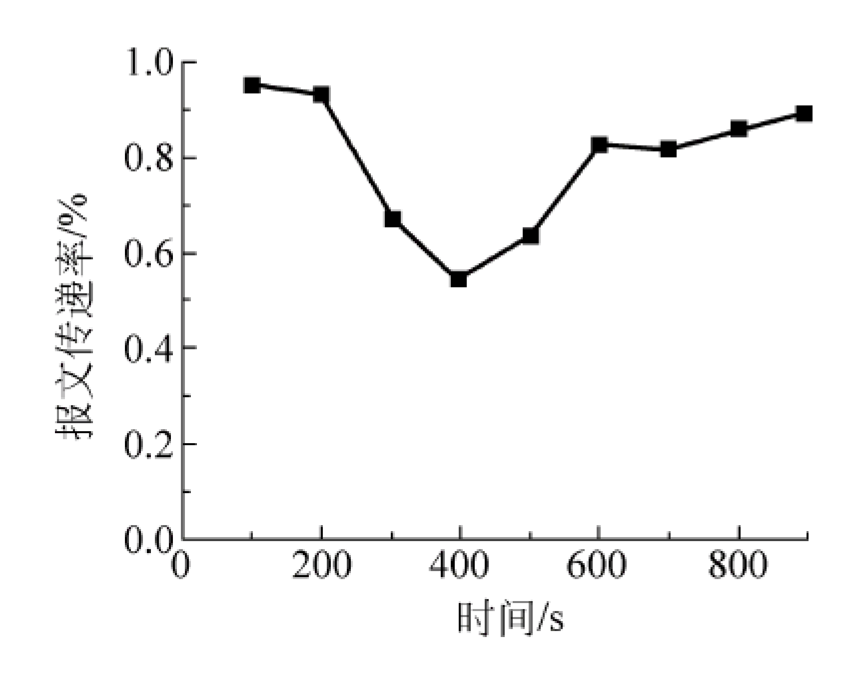


图6-7 入侵响应前后分组传递率变化图

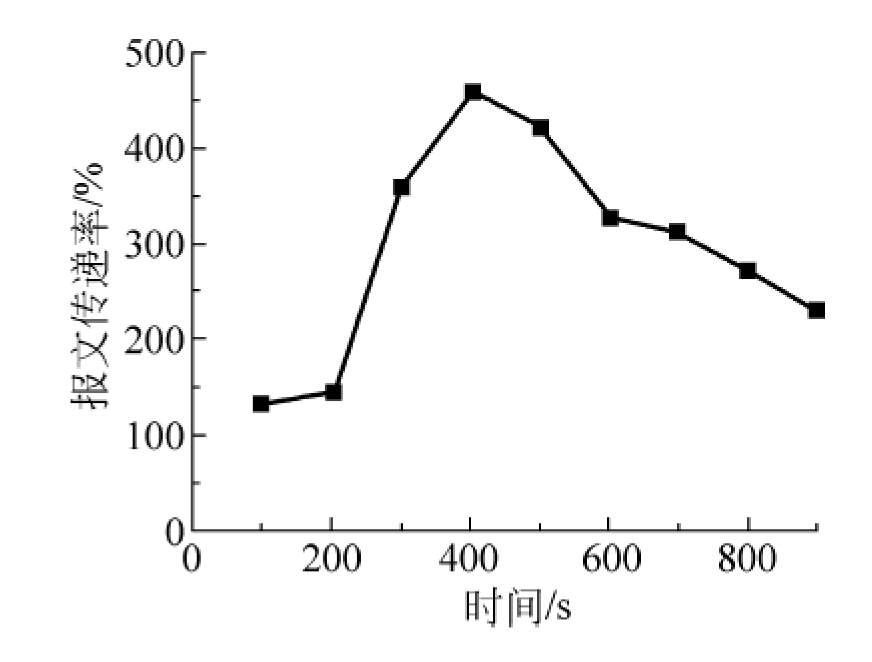


图6-8 入侵响应前后分组传输延迟变化图

6.7 小结

本节介绍一种基于移动代理的主动入侵响应模型，通过监视代理实现对整个无线自组织网络中每个节点行为的监视，将各个节点的行为信息送往决策代理进行判断，识别入侵者后，通过阻击代理形成移动防火墙将入侵者包围并隔离，最终阻止了入侵行为。模拟实验证实主动入侵响应模型的有效性。

参考文献

[1] Fred Cohen. Simulating Cyber Attacks, Defenses, and Consequences, http://all. net/journal/ntb/simulate/simulate. html.

[2] Hu Y C. Perrig A. A Survey of Secure Wireless Ad hoc Routing, IEEE Security & Privacy Magazine, May-June 2004, 2(3): 28-39.

[3] Khalili A, Katz J, Arbaugh W. Towards Secure Key Distribution in Truly Ad-hoc Networks IEEE Workshop on Security and Assurance in Ad hoc Networks, in conjunction with the 2003 International Symposium on Applications and the Internet, Orlando, FL, January 28, 2003.

[4] Kuang X H, Zhu P D, Lu X C. Distributed Group Rekeying Algorithms for Mobile Ad-hoc Networks. Journal of Software, 2004, 15(5): 757-766.

[5] Yan X, Miao F Y, Zhang W C, et al, Secure Distributed Authentication Ba sed on Multi2Hop Signing with Encrypted Signature Functions in Mobile Ad hoc Networks, ACTA ELECTRONICA SINICA, 2003, 2.

[6] Zhou, Haas Z J. Securing Ad hoc Networks. IEEE Networks Special Issue on Network Security, November/December, 1999.

[7] Srdjan Capkun, Levente Nuttyan, Jean-Pierre Hubaux. Self-Organized Public-key Management for Mobile Ad hoc Networks. IEEE Transactions on Mobile Computing, 2003, 2(1).

[8] Papadimitratos P, Haas Z. Secure routing for mobile Ad hoc Networks, in Proceedings of the SCS

communication Networks and Distributed Systems Modeling and Simulation Conference, San Antonio, TX, January, 2002: 27-31.

[9] Hu Yih Chun, Adrian Perrig, Johnson D B. Ariadne: A secure On-Demand Routing Protocol for Ad hoc Networks, in Proceedings of the MobiCom 2002, September 23-28, 2002, Atlanta, Georgia, USA.

[10] Kimaya Sanzgiri, Bridget Dahill, Brian Neil Levine, et al. A Secure Routing Protocol for Ad hoc Networks, In Proceedings of 2002 IEEE International Conference on Network Protocols (ICNP), November 2002.

[11] Hu Yih Chun, Johnson D B, Adrian Perrig. SEAD: Secure Efficient Distance Vector Routing for Mobile Wireless Ad hoc Networks, in Proceedings of the 4th IEEE Workshop on Mobile Computing Systems & Applications (WMCSA 2002), pp. 3-13, IEEE, Calicoon, NY, June 2002.

[12] Zhang Yongguang, Lee Wenke. Intrusion Detection Techniques for Mobile Wireless Networks, Mobile Networks and Applications, 2003.

[13] Farooq Anjum, Dhanant Subhadrabandhu and Saswati Sarkar Signature based Intrusion Detection for Wireless Ad-hoc Networks: A Comparative study of various routing protocols Proceedings of Vehicular Technology Conference, Wireless Security Symposium, Orlando, Florida, October 2003.

[14] Zhang Yongguang, Lee Wenke, Intrusion Detection in Wireless Ad-hoc Networks Proceedings of The Sixth International Conference on Mobile Computing and Networking (MobiCom 2000), Boston, MA, August 2000.

[15] Schnackenberg D, Holliday H, Smith R, et al. Cooperative Intrusion Traceback and Response Architecture (CITRA), Proceedings of the Second DARPA Information Survivability Conference and Exposition (DISCEXII), Anaheim, CA, June 2001.

[16] Schnackenberg D, Djahandari K, Sterne D. Infrastructure for Intrusion Detection and Response, Proceedings of the DARPA Information Survivability Conference and Exposition, Hilton Head, SC, January 2000.

[17] Dan Sterne, Kelly Djahandari, Brett Wilson, et al. Autonomic Response to Distributed Denial of Service Attacks, In Proceedings of the 4th International Symposium on Recent Advances in Intrusion Detection, RAID 2001, Davis, CA, USA, October 2001, Springer-Verlag, pp. 134-149.

[18] Carver C, Hill J, Surdu J, et al. A methodology for using Intelligent Agents to provide Automated Intrusion Response, Proceeding of the IEEE Workshop on Information Assurance and security, United States Military Academy, West Point, USA, June 2000.

[19] Tennenhouse D L. A Survey of Active Network Research. IEEE Communications Magazine, 1997, 35(1): 80-86.

[20] Wang X, Reeves D, Wu S F, et al. Sleepy Watermark Tracing: An Active Intrusion Response Framework. the Proceedings of 16th International Conference of Information Security, Paris, France, 2001.

[21] Wang X, Reeves D, Wu S F. Tracing-Based Active Intrusion Response, Journal of Information Warfare, 2001, 1(1).

第7章　无线局域网的安全

摘要：无线局域网(WLAN)是近年来发展迅速的无线数据通信网，但在发展同时，它又面临着许多安全问题。本章首先对无线局域网进行了概述，然后对无线局域网的安全风险和安全需求进行了分析，最后重点阐述了无线局域网的安全技术以及安全协议。

关键字：无线局域网、安全风险、安全需求、安全技术、安全协议。

7.1　概述

无线局域网(Wireless Local Area Network，WLAN)是高速发展的现代无线通信技术在计算机网络中的应用，它采用无线多址信道的有效方式支持计算机之间的通信，并为通信的移动化、个人化和多媒体应用提供了实现的手段。随着个人数据通信的发展，功能强大的便携式数据终端以及多媒体终端得到了广泛应用。为了实现任何人在任何时间、任何地点均能进行数据通信的目标，要求传统的计算机网络由有线向无线、由同定向移动、由单一业务向多媒体发展，顺应这一需求的无线局域网技术因此得到了普遍的关注。无线局域网以其方便、快捷、廉价等诸多优势，在企事业内部和公共热点地区等领域的应用中很快取得了长足的发展和巨大的成功。

7.1.1　无线局域网协议栈

IEEE 802.11 定义了无线局域网物理层和介质访问控制子层 MAC 层规范。其协议栈如图 7-1 所示，对应于 OSI 模型中的物理层和数据链路层中的 MAC 子层。其中物理层定义了通过无线连接所必需的机械和电气特性；媒体访问控制层则定义了两个数据链路层之间建立和维持数据传输，并将数据流无差错的提供给网络层的功能。

1. 物理层关键技术

在较为复杂的电磁环境中，多径效应、频率选择性衰落和其他干扰源的存在使得实现无线信道的高速数据传输比有线信道更困难，WLAN 需要采用合适的调制技术。

(1) 扩频。所谓扩频通信，简单地说，它是一种信息传输方式，其信号所

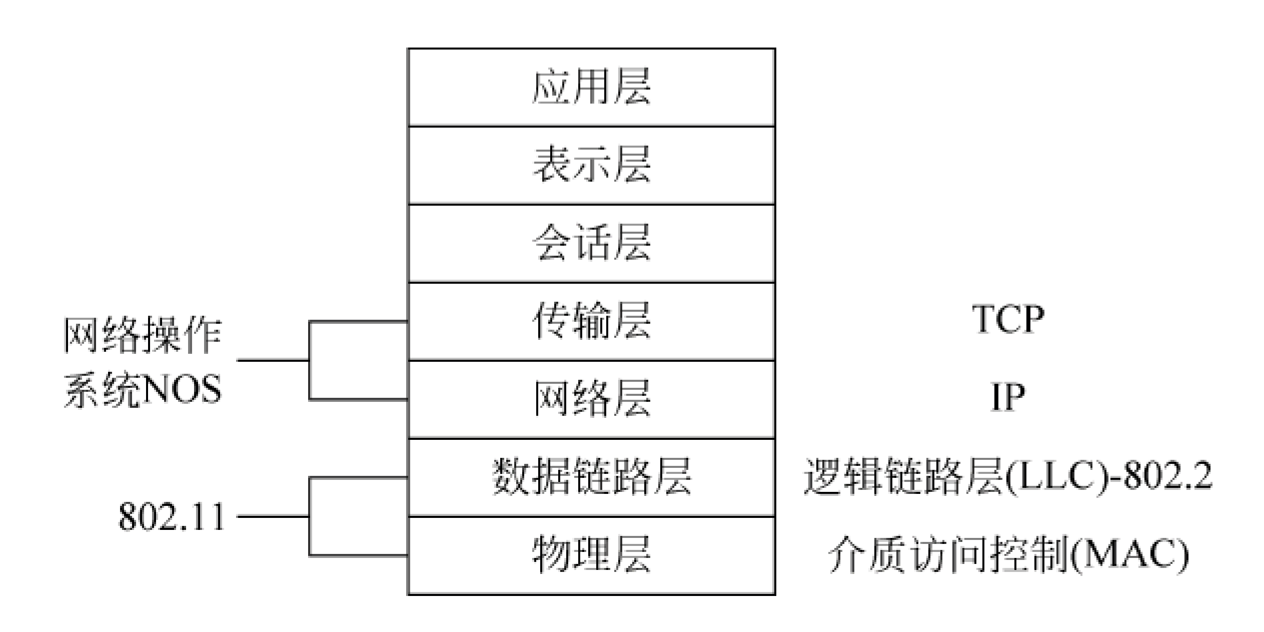


图 7-1　802.11 与 OSI 模型

占有的频带宽度远大于所传信息必需的最小带宽；频带的展宽是通过编码及调制的方法实现的，并与所传信息数据无关；在接收端则用相同的扩频码进行相关解调来解扩及恢复所传信息数据，具有较高抗干扰能力和较强的保密性。扩频技术包括以下几种方式：直接序列扩展频谱，简称直扩(DS)、跳频(FH)、跳时(TH)、线性调频(chirp)。此外，还有这些扩频方式的组合方式，如 FH/DS、TH/DS、FH/TH 等，在通信中应用较多的主要是 DS、FH 和 FH/DS。

(2) DSSS 调制技术。基于 DSSS 的调制技术有三种。最初 IEEE 802.11 标准制定在 1Mb/s 数据速率下采用 DBPSK。如提供 2Mb/s 的数据速率，要采用 DQPSK，这种方法每次处理两个比特码元，成为双比特。第三种是基于 CCK 的 QPSK，是 IEEE 802.11b 标准采用的基本数据调制方式。它采用了补码序列与直序列扩频技术，是一种单载波调制技术，通过 PSK 方式传输数据，传输速率分为 1Mb/s、5.5Mb/s 和 11Mb/s。CCK 通过与接收端的 Rake 接收机配合使用，能够在高效率的传输数据的同时有效地克服多径效应。IEEE 802.11b 使用了 CCK 调制技术来提高数据传输速率，最高可达 11Mb/s。但是传输速率超过 11Mb/s，CCK 为了对抗多径干扰，需要更复杂的均衡及调制，实现起来非常困难。因此，802.11 工作组，为了推动无线局域网的发展，又引入新的调制技术。

(3) PBCC 调制技术。PBCC(Packet Binary Convolution Coding)调制技术是由 TI 公司提出的，已作为 802.11g 的可选项被采纳。PBCC 也是单载波调制，但它与 CCK 不同，它使用了更多复杂的信号星座图。PBCC 采用 8PSK，而 CCK 使用 BPSK/QPSK；另外 PBCC 使用了卷积码，而 CCK 使用区块码。因此，它们的解调过程是十分不同的。PBCC 可以完成更高速率的数据传输，其传输速率为 11Mb/s、22Mb/s 和 33Mb/s。

(4) OFDM 技术。OFDM 技术是一种无线环境下的高速多载波传输技术。无线信道的频率响应曲线大多是非平坦的，而 OFDM 技术的主要思想：就是在频域内将给定信道分成许多正交子信道，在每个子信道上使用一个子载波进行调制，并且各子载波并行传输，从而有效的抑制无线信道的时间弥散所带来的 ISI。由于在 OFDM 系统中各个子信道的载波相互正交，于是它们的频谱是相互重叠的，这样不但减小了子载波间的相互干扰，同时又提高了频谱利用率。

由于无线信道存在频率选择性，所有的子信道不会同时处于比较深的衰落情况中，因此可以通过动态比特分配以及动态子信道分配的方法，充分利用信噪比高的子信道，从而提升系统性能。由于窄带干扰只能影响一小部分子载波，因此 OFDM 系统在某种程度上抵抗这

种干扰，OFDM 系统结构如图 7-2 所示。IEEE 802.11 a/g 标准为了支持高速数据传输都采用了 OFDM 调制技术。目前，OFDM 结合时空编码、分集、干扰（包括符号间干扰 ISI 和邻道干扰 ICI）抑制以及智能天线技术，最大限度的提高物理层的可靠性。如再结合自适应调制、自适应编码以及动态子载波分配、动态比特分配算法等技术，可以使其性能进一步优化。

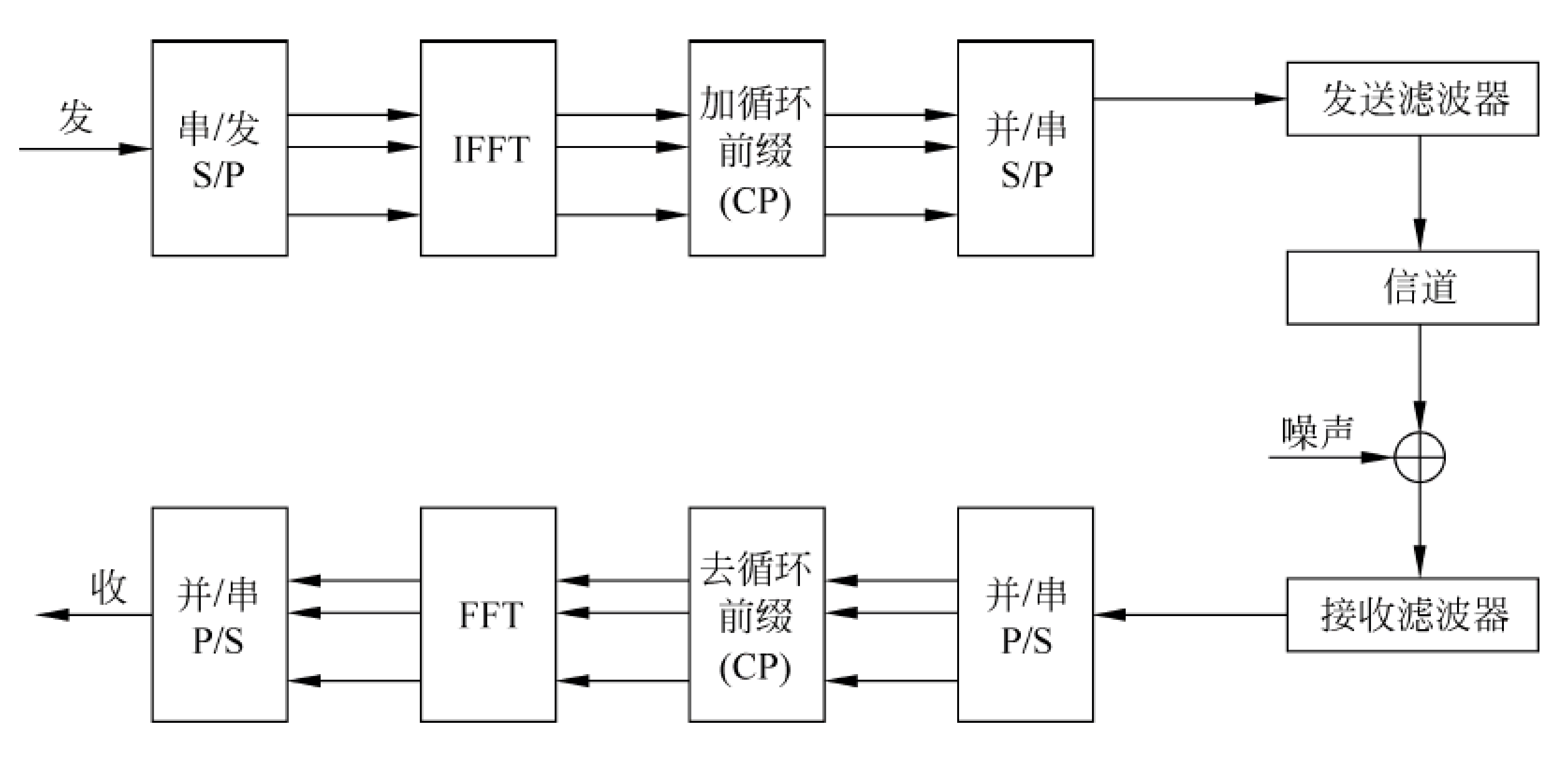


图 7-2　OFDM 系统结构框架图

（5）MIMO OFDM 技术。MIMO 技术能在不增加带宽的情况下成倍地提高通信系统的容量和频谱利用率。它可以定义为发送端和接收端之间存在多个独立信道，也就是说天线单元之间存在充分的间隔，因此消除了天线间信号的相关性，提高信号的链路性能增加了数据吞吐量。现代信息论表明：对于发射天线数为 N，接收天线数为 M 的多入多出（MIMO）系统，假定道信为瑞利衰落信道，并设 N、M 很大，则信道容量 C 近似为公式 $C=[\min(M,N)B\log_2(\rho/2)]$（其中 B 为信号带宽，ρ 为接收端平均信噪比，$\min(M,N)$ 为 M、N 中的较小者）。这表明，MIMO 技术能在不增加带宽的情况下成倍地提高通信系统的容量和频谱利用率。研究表明，在瑞利衰落信道环境下，OFDM 系统非常适合使用 MIMO 技术来提高容量。采用 MIMO 系统是提高频谱效率的有效方法。我们知道，多径衰落是影响通信质量的主要因素，但 MINO 系统却能有效地利用多径的影响来提高系统容量。系统容量是干扰受限的，不能通过增加发射功率来提高系统容量，而采用 MIMO 结构不需要增加发射功率就能获得很高的系统容量。因此将 MIMO 技术与 OFDM 技术相结合是下一代无线局域网发展的趋势。

2. MAC 子层技术

由于在无线网络中冲突检测较困难，IEEE 802.11 规定介质访问控制（Medium Access Control，MAC）子层采用冲突避免（Collision Avoid，CA）协议，而不是冲突检测（Collision Detect，CD）协议。为了尽量减少数据的传输碰撞和重试发送，防止各站点无序地争用信道，无线局域网中采用了与以太网 CSMA/CD 相类似的 CSMA/CA（载波监听多路访问/冲突防止）协议。CSMA/CA 通信方式将时间域的划分与帧格式紧密联系起来，保证某一时刻只有一个站点发送，实现了网络系统的集中控制。因传输介质不同，CSMA/CD 与 CSMA/CA 的检测方式也不同。CSMA/CD 通过电缆中电压的变化来检测，当数据发生碰撞时，电缆中的电压就会随着发生变化；而 CSMA/CA 采用能量检测（Energy Detect，ED）、

载波检测(Carrier Sense,CS)和能量载波混合检测三种检测信道空闲的方式。

IEEE 802.11 定义基本服务群(Base Service Set,BSS)是无线局域网的基本单元,它的功能包括分布式协调功能(Distributed Coordination Function,DCF)和无线访问接入点协调功能(Point Coordination Function,PCF)。协调功能是决定在 BSS 内工作的一个站,通过无线介质何时允许发送和可能接收协议单元的逻辑功能。DCF 是 IEEE 802.11 标准中 MAC 协议的基本介质访问方法,它作用于基本服务群和基本网络结构中,可在所有站实现,它支持竞争型异步业务,而 PCF 可支持无竞争型时限业务及无竞争型异步业务。在 PCF 模式中,将由一个无线访问接入点(Access Point,AP)来控制对介质的所有访问。在系统处于 PCF 模式期间,负责控制访问的无线访问接入点将接纳每个端站的数据,经过给定的时间后转移到下一站。除非某个端站被接纳,否则它不允许进行发射。也只有当端站被接纳以后,它们才接收来自无线访问接入点的数据。由于 PCF 按预定的方式给每个端站确定了一个发射顺序,因此保证了每条数据流的最大延迟。PCF 的不足之处是它的可伸缩性较差,因为是由单一的无线访问接入点来控制媒体访问,且必须接纳所有端站的通信,所以这种做法在较大的网络中效率较低。

IEEE 802.11 的 MAC 层负责客户端与无线访问接入点之间的通信。当一个 802.11 客户端进入一个或多个无线访问接入点的覆盖范围时,它将根据信号强度和监测到的包错误率,选择其中性能最好的一个无线访问接入点并与之联系。一旦被该无线访问接入点接受,客户端会将无线信道调整到设置无线访问接入点的无线信道。它定期检测所有的 802.11 信道,以便确定是否有其他的无线访问接入点能够提供更好的性能。如果检测到存在这样一个无线访问接入点,它将与新的无线访问接入点重新建立联系,客户端将调整到设置该无线访问接入点的无线信道。出现这样的重新连接通常是由于无线端站在物理位置上离开了原始无线访问接入点,导致信号变弱。此外,当建筑物中的无线特性发生变化,或者原始无线访问接入点的网络通信量过高时,也会出现重新联系的情况。在后一种情况下,这个功能一般称为“负载平衡”,因为它的主要作用是将总体的无线 LAN 负载最有效地分布到可用的无线基础设施中。

IEEE 802.11 中 MAC 提供的服务有:安全服务、MSDU 重新排序服务和数据服务。其中安全服务提供的服务范围局限于站与站之间的数据交换,其内容为:加密、验证、与层管理实体相联系的访问控制。IEEE 802.11 标准中提供了 WEP(Wired Equivalent Privacy)加密算法,其目标是为无线 LAN 提供与有线网络相同级别的安全保护。另外,为了进行访问控制,ESSID 将被放置到每个无线访问接入点中,它是无线客户端与无线访问接入点联系所必不可少的。除此之外,每个无线访问接入点中还包括一个有关 MAC 地址的访问控制列表,只有那些 MAC 地址在这个列表中的客户端才能对无线访问接入点进行访问。MSDU 重新排序服务是为了提高成功发送的可能性,只有在节电方式工作下的站,且不处于激活状态,才可先将 MSDU 缓存起来,等站激活时再突发出去,对缓存数据进行重新排序。MAC 数据服务可使对等 LLC 实体进行数据单元的交换,本地 MAC 利用下层的服务将一个 MSDU 传给一个对等的 MAC 实体,然后又传给对等的 LLC 实体。

IEEE 802.11 物理层的无线媒体决定了 WLAN 具有独特的 MAC 机制。IEEE 802.11 支持两种不同的 MAC 方案:第一种方案是分布协调功能(Distributed Coordination Function,DCF),其中 DCF 采用了载波多路访问/冲突避免技术(CSMA/CA),类似于以太

网标准中的CSMA/CD,支持异步数据传输等异步业务,所有要传输数据的用户拥有平等接入网络的机会。DCF基于载波侦听多址接入/碰撞预防(CSMA/CA)而未采用有线LAN中主要使用的载波侦听多址接入/碰撞检测(CSMA/CD)多址接入方式。这是因为站点传输时听不到信道碰撞。由于无线信道动态范围大,在有效带宽内采用碰撞检测方式是很困难的,故只能采取随机退避方式以减少两帧碰撞的概率。第二种方案是点协调功能(Point Coordination Function,PCF),是DCF方式的一个补充。在这种方式下,由AP控制移动站点的数据发送,不存在信道竞争的问题。基于由接入点控制的轮询(poll)方式,主要用于传输实时业务。MAC子层由DCF和PCF两部分组成。DCF直接位于物理层之上。所有站点均支持DCF。在Ad hoc网中,DCF独立工作;在基本结构网中,DCF可独立工作也可与PCF共同工作。MAC子层负责信道分配过程、PDU寻址、帧形成、差错校验、分组拆装。传输媒体可工作于竞争方式,每个站点传输任意一分组时需对信道进行竞争接入;媒体也可在竞争期(Contention Period,CP)和非竞争期(Contention Free Period,CFP)间交替工作。在非竞争期间,媒体的使用由接入点控制或作为中介,因而站点不必进行接入信道竞争。

7.1.2　无线局域网组成

802.11无线局域网主要由无线网卡和接入点AP等网络设备构成。其中无线网卡主要用来收发无线局域网的数据报文,而AP则是主要为了实现有线局域网和无线局域网之间的互联和通信,它有两个功能:桥接功能和移动管理。桥接功能是指在有线/无线局域网中数据包的存储、转发和过滤等;移动管理功能主要包括对移动端的认证、登录和散步的管理等。

802.11无线局域网按照其功能可以划分为三个部分:基于无线媒体的通信网络、提供综合业务的用户终端以及支持网络运行的管理单元。通信网络包括基本服务区(Base Server Set,BSS)、扩展服务区(Extend Service Set,ESS)以及无线接入点(AP)等。BSS是指由无线收发机及地理环境所确定的通信覆盖区域,通常称之为小区(Cell)。无线接入点为分散的BSS提供中远距离的点对点连接,从而构成扩展服务区。用户终端可以是台式计算机、便携式计算机等设备。网络管理单元由网络的整体配置和各主要模块(设备、软件)配置组成,它包括协议转换、网络寻址、路由选择、速率匹配、差错控制、密钥管理等。

7.1.3　无线局域网的拓扑结构

802.11无线局域网支持2种网络拓扑结构:基于AP的网络结构和基于P2P的网络结构。

1. 基于AP的网络结构

基于AP的网络结构也称有中心网络结构。它由无线AP、无线工作站STA(Station)以及DSS构成,覆盖的区域分BSS和ESS。无线访问点也称无线AP或无线hub,用于在无线网络和有线网络之间接收、缓存和转发数据。所有的移动终端之间,移动终端与有线网络中的主机之间的通信都通过AP。该结构是有线网络向无线网络的扩展,可以通过放置多个AP来扩展无线覆盖范围,并允许移动终端在不同AP之间漫游。无线AP通常能够覆盖几十至几百用户,覆盖半径达上百米,如图7-3所示。

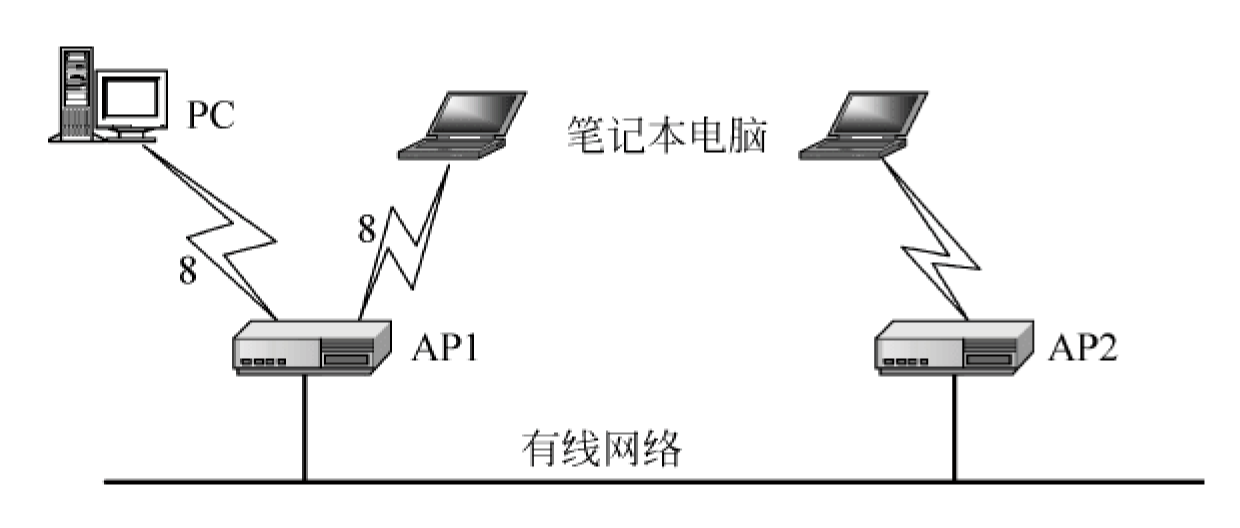


图 7-3　基于 AP 的网络结构

BSS 由一个无线访问点以及与其关联的无线工作站构成，在任何时候，任何无线工作站都与该无线访问点关联。换句话说，一个无线访问点所覆盖的微蜂窝区域就是基本服务区。接入点类似于蜂窝通信网中的基站，可以看做是将 IEEE 802.11 网连到有线骨干网的网桥。接入点通过提供多个基本业务群互连的连接点来扩展通信范围，形成扩展业务群，如图 7-4 所示。

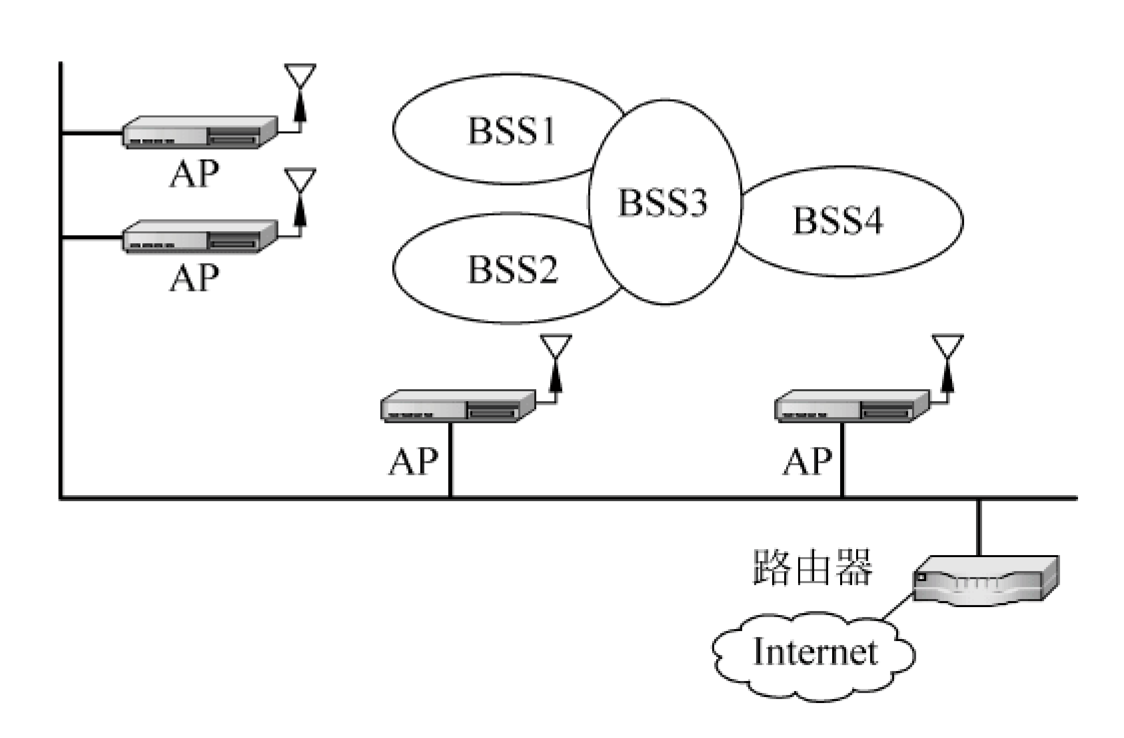


图 7-4　ESS 网络结构

扩展业务群(Extend Service Set，ESS)包括多个基本业务群，通过 DS 连到一起。DS 可以是 IEEE 802.3 以太网、IEEE 802.4 令牌总线或其他 IEEE 802.11 无线媒体。扩展业务群也可通过入口部件 Portal 为无线用户提供至有线网(如 Internet)的网关接入。Portal 是规定 IEEE 802.11 网与非 IEEE 802.11 网在 DS 上的连接点的逻辑实体。如果网络 IEEE 802.X，Portal 功能类似于网桥。要求一个无线站点充当中心站，所有站点对网络的访问均由其控制。

由于每个站点只需在中心站覆盖范围之内就可与其他站点通信，故网络中心站布局受环境限制亦小。此外，中心站为接入有线主干网提供了一个逻辑接入点。有中心网络拓扑结构的弱点是抗毁性差，中心点的故障容易导致整个网络瘫痪，并且中心站点的引入增加了网络成本。在实际应用中，无线网一般与有线主干网络结合起来使用。这时，中心站点充当无线网与有线主干网的转接器。

2. 无中心网络结构

无中心网络也称对等网络(Peer to Peer)、无 AP 网络或 Ad hoc 网络。这种模式下网络不需要 AP 就可以实现，移动站点通过无线接口点对点的直接通信，无中心拓扑的网络要求网中任意两个站点均可直接通信，该网络无法接入有线网络中，只能独立使用。

采用这种拓扑结构的网络一般是用公用广播信道，各站点都可竞争公用信道，而信道接入控制(MAC)协议大多采用 CSMA(载波监测多址接入)类型的多址接入协议。这是最简单的无线局域网结构，如图 7-5 所示。

一个对等网络由一组有无线接口的计算机组成。这些计算机要有相同的工作组名、ESSID 和密码。对等网络组网灵活，任何时间只要两个或更多的无线接口互相都在彼此的范围之内，它们就可以建立一个独立的网络。这些根据要求建立起来的典型网络在管理和预先调协方面没有任何要求。这种结构的优点是网络抗毁性好、建网容易、且费用较低。但当网中用户数(站点数)过多时，信道竞争成为限制网络性能的要害。并且为了满足任意两个站点可直接通信，网络中站点布局受环境限制较大。Ad hoc 网由某基本业务群(BSS)内的一组站点组成，又称为 IBSS，基本业务群是 IEEE 802.11 结构的基本功能块，它覆盖的地理区域类似于蜂窝通信网中的蜂窝。在基本业务群内，任何一个站点可与其他任何一个站点直接建立通信过程。

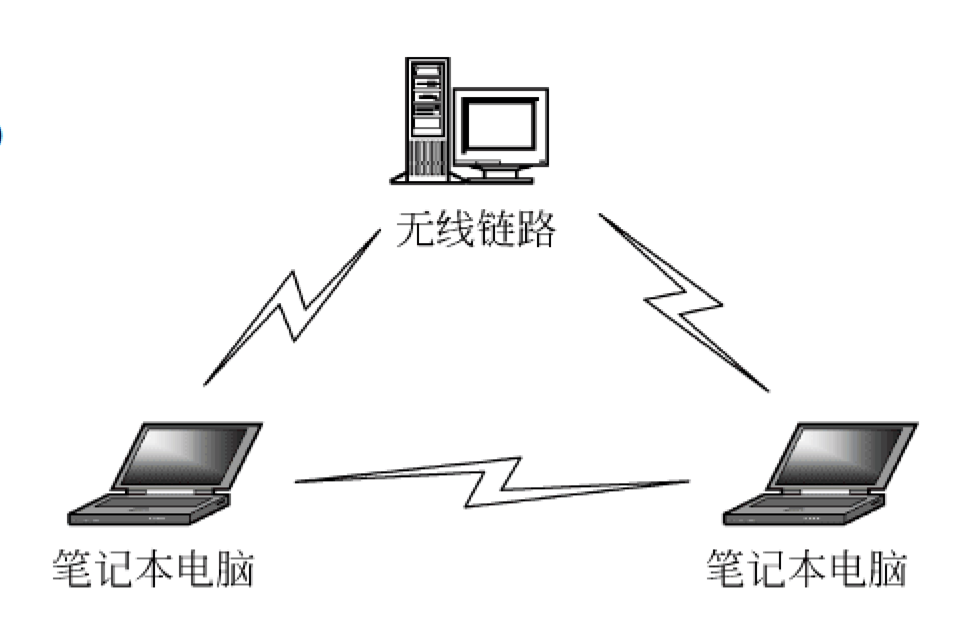


图 7-5　无中心网络结构

7.1.4　无线局域网的应用及发展趋势

IEEE 802.11 最初只是作为一种无线接入协议，而问世后可谓是异军突起，目前，Wi-Fi 技术已经被认为是无线宽带发展的新方向。在美国，像 Nextel、Cingular 这样的移动运营商正在商业楼宇中部署 Wi-Fi 网络，Bellsouth、Verizon 等固定运营商也不甘落后。Verizon 已经在纽约启动了 150 个热点地区的 Wi-Fi 网络，并在纽约部署 1000 多个 Wi-Fi 网络。2005 年底，由 Intel、IBM 和 AT&T 合作组建的 Commeta 网络公司，更是定下了在全美 50 个大城市建设 2 万个热点 Wi-Fi 网络的宏伟计划。利用 Wi-Fi 802.11g 将传输速率提升到 54Mb/s，美国 Vivato 公司推出的一款新型交换机能把目前 Wi-Fi 无线网络 100m 的通信半径扩大到 6.5km，同时用户接纳数量大幅度增加。诺基亚和摩托罗拉等公司也加入到 WLAN 的研发与推广之中。Wi-Fi 网络在家庭网络中越来越普及了，并首次超过了以太网。据市场调研公司 Parks Associates 进行的调查显示，在部署有网络的家庭中，52%的家庭使用了无线网络技术，而使用以太网和电力线网络技术的比例分别为 50%和约 5%(三者相加超过 100%的原因是一些家庭采用了一种以上的网络技术)。

在欧洲，引领 Wi-Fi 的运营商是英国电信和瑞士电信。这些没有受 3G 牌照拖累的企业，大展拳脚，期望通过投资 Wi-Fi，花少得多的资金在移动数据市场挤占一定的份额。德国的 T-mobile 公司也对 Wi-Fi 情有独钟，该公司已与 1400 多家星巴克咖啡店联合建设了 Wi-Fi 网络，并计划在全球 2000 多个社区提供 Wi-Fi 接入服务。

国内的电信运营商正在以 FTTX＋WLAN、ADSL＋WLAN 和 GPRS＋WLAN、CDMA 1X＋WLAN 等形式进军 Wi-Fi 领域。这股风潮势必将牵动无线局域网产业链中的各个环节。作为国内最早涉足 Wi-Fi 领域的运营商，中国网通已经在无线局域网布点 1000 多个，大多数都集中在商务客人经常出入的热点地区，如机场、商务酒店、会展中心等，其“无限伴旅”(Mobile Office)的无线局域网接入服务已经在北京、上海、广州、深圳等城市展开。

随着 Wi-Fi 热潮在全球的兴起,国内其他的运营商也迅速跟进,加入这一市场的争夺。与网通紧盯商务用户不同,中国电信充分利用其有线资源,将 WLAN 与 ADSL 捆绑,迅速夺取了国内最大的市场份额。

中国电信名为“天翼通”的 Wi-Fi 无线宽带接入业务已经在上海和广东等地铺开,目标直指普通消费者。而开始还在犹豫的中国移动和中国联通也迅速跟进。2005 年世界电信日宣布进军 Wi-Fi 市场的中国移动,第一期动用 18 亿元在全国 32 个城市推广 WLAN 业务,与 GPRS 进行捆绑。2006 年 5 月 17 日,中国移动宣布已在全国近 700 个机场、酒店等热点地区实现了 WLAN 覆盖。另外,联通也正在考虑把 CDMA 1X 与 WLAN 捆绑。

当人们广泛期待的 3G 时代并未如期而至,无线局域网的高速发展更是成为电信业发展的一大亮点。全球无线局域网销售收入在 2004 年为 100 亿美元,预测到 2008 年将增长到 440 亿美元,年增长率为 44%。由此可见,基于 IEEE 802.11 技术的 WLAN 已经成为目前宽带无线网络接入技术的主流,尤其是随着 802.11 系列规范的相继出台,未来的 802.11 必将会重现 802.3 的辉煌历史,未来的无线网络终端接入技术将会进入 802.11 系列技术时代。

由于下一代无线网络将是由三个部分组成,即无线接入网、核心网和骨干网三部分组成的。各种移动网和无线网都采用 IP 技术成为互联网的无线接入网络,移动或无线终端就可以通过无线方式接入互联网,享受互联网信息服务,并在互联网平台上进行通信。根据 2006 年中国下一代互联网示范工程产业化及应用试验专项规划,重点进行 WWAN 和 WLAN 的互通融合技术和业务示范,研究未来无线局域网中各种通信系统的互通、融合技术及其网络构架,建设基于移动 IPv6 的 WWAN 和 WLAN 融合网络,开展业务试验和应用示范。

3G 系统旨在提供一个全覆盖、高质量保证的通信网络,可以提供语音和数据业务,数据传输速率达 2Mb/s。从目前的移动应用业务角度来看,传输速率最高为 2Mb/s 显然无法满足用户对高速传输速率的需求。另外,由于 3G 系统所使用的频率为 2GHz 频段,这是非常珍贵和短缺的频率资源,运营商为了获得 3G 牌照需要花费大量资金;另一方面,3G 网络的单基站覆盖范围 1.5~8km,要达到全网覆盖,需要设置大量的网络设备,造成运营成本非常高。

WiMAX[11] 移动通信系统主要定位于分组数据的业务传输,其峰值数据的速率可达到 75Mb/s,比 3G 系统高得多,但其主要为固定、便携或低速移动的用户提供接入,在网络建设初期和中期阶段并不支持高速移动下的无缝漫游。而 3G 移动通信系统具有支持快速漫游以及提供全网覆盖的通话业务功能优点。

WLAN 作为 WiMAX 网络的补充技术,满足局部热点地区提供较高数据传输,为解决该问题提供了一种新的途径。最新的 802.11s 草案,引入了 Mesh 机制,Mesh 网络具有自愈和自配置的优越性。研究成果表明,未来的 WLAN 组网将是采用 Mesh 网络技术的混合网络体制。

7.2 安全风险与安全需求

现在,无线网络已经开始进入人们的日常生活之中。目前存在着三种代表性的无线网络技术,这三种技术都具有各自的优势和特点。在广域网融合方面,蜂窝无线技术风头正

劲；在局域网方面，无线局域网(IEEE 802.11 标准)已经投入应用；个人区域网络则主要采用蓝牙技术。一旦用户做出自己的网络将支持哪个无线平台的决定，剩下的最主要问题就是如何保证系统的安全了。

7.2.1 无线局域网的安全风险分析

无论是有线网络还是无线网络都受到安全问题的困扰，只要内部网的合法用户收发电子邮件、访问内部网或互联网系统，就已经将内部网暴露于一个范围广泛的潜在威胁之中。无线设备同样也可以将内部网暴露于一个潜在的安全威胁之中，这些设备很容易由于使用人员的疏忽而发生被盗，使得秘密信息落入竞争者或敌对者的手中[10]。

近年来，各企业或机构已经越来越重视利用防火墙保护其内部网以防受到来自外部网的攻击。但是，无线网络打开了一个允许攻击者超越企业或机构的物理安全界线而进入内部网的后门。攻击者可以在该企业或机构的停车场里访问企业或机构内部网的主机，这就是所谓的“停车场”攻击。值得注意的是，在一些情况下，防火墙使得内部网在攻击者面前显得更加脆弱，这是由于不恰当的配置会使系统更易受到攻击和潜在的威胁。为了便于用户使用，网络系统总是尽量做到开放和易于互联，但这同时又给网络系统带来了安全隐患。无线网络除了受到与有线网络相同的安全威胁以外还有其特殊性。

1. 无线局域网所面临的安全风险

1）来自网络外部或内部的窃听

无线传输的介质是共享的，也正是由于这个原因，相对于有线网络来说，通过无线局域网发送和接收数据时就更容易被窃听。目前的无线局域网使用 2.4GHz 范围的无线电波进行网络通信。任何人都可以用一台带无线网卡的 PC 或者廉价的无线扫描器进行窃听。而且无线局域网易于被发现，但严格来说这并不是安全隐患。因为所有的无线网络都要通告它的存在以便用户能够与之建立连接并使用它提供的网络服务。

IEEE 802.11 标准需要网络周期性地利用特殊的帧通告它的存在，该帧称为信标帧。然而，加入网络所需的信息同样也是攻击网络所需的信息。标信帧没有经过任何加密处理，这意味着一个 IEEE 802.11 网络及其参数可以被任何拥有 IEEE 802.11 卡的人利用。黑客工具 War driver 可以使用高性能天线和相应的软件来登录信标帧并用 GPS 与其保持连接。因此，当运行无线局域网时，网络管理员所要面对的最大问题之一就是数据安全问题，也就是防止数据被窃听。这种窃听可能来自外部，也可能来自内部，数据加密是防止这类入侵的最有效方法。

在有线环境里，对物理线路接入的限制可以使他人只能在网络所在的建筑物外徘徊，而不能连接到内部网络里。但是在无线局域网环境中，对 AP 来说，它无法知道是否有操纵无线设备的人在网络所在的建筑物里。目前还无法解决短程移动通信进入隔离严密的办公空间而不泄露无线信号这个问题。减轻非法访问的风险可以利用强有力的访问控制和加密技术，这样就可以防止无线局域网成为网络中易于登录的点。配置 AP 外的防火墙并且利用 VPN 来保护敏感的通信信息也是解决该风险的一种方法。

IEEE 802.11 不提供防护被动观察通信流量攻击的方法。对此来说主要风险就是 IEEE 802.11 不提供在面对窃听时保护数据安全的方法。对于无线网络分析者来说，帧头

总是“清晰可见的”。很多对 WEP 协议有异议的人认为不受窃听的侵害这一点应当得到保证。尽管目前符合 IEEE 802.11b 标准的 AP 和网卡支持最基本的安全措施，但大多数企业或机构甚至还没有启用它们。而那些采取了安全措施的企业或机构会在安全问题上产生错觉，认为 WEP 这一脆弱的协议以及密钥管理系统所能提供的保护并没有多大价值。实际上。由于无线局域网被设计成只能覆盖有限的区域，所以现行的 WEP 安全级别也许是足够的。但是，对于需要更高级别保护的企业或机构来说就远远不够了。

2）来自黑客的攻击

无线局域网的快速发展使得安全问题成为急需解决的问题。对黑客来说，无线局域网规模大了他们实施攻击的机会也就多了。而且随着技术的进步，目前黑客在实施攻击时不一定再需要有非常昂贵的硬件和高超的技术，而只需要一台计算机和一个免费软件就可以了。因此，现在黑客攻击事件也就越来越多了。黑客凭借 War driver 技术可以攻击无线局域网，他们通过携带笔记本电脑和 IEEE 802.11 以太网卡能够轻易地找到未加保护的无线局域网，随之可以偷窃数据资料和施放病毒，也可能会蓄意破坏网页或者从网络内部发动匿名攻击。这种新的黑客行为日益嚣张，它伴随着无线局域网的壮大而发展。其中的主要原因是很多无线局域网完全暴露在黑客面前，没有任何自卫能力。所以，无线局域网用户人数的不断增长，也就为黑客攻击无线网络提供了更多的机会。为了安全，最好的办法就是把无线局域网“锁”起来，使其避免遭受攻击。但是，现有的技术水平要做到这一点还不太现实。目前比较现实的方法只能是：建立安全的程序并且利用加密手段来确保一定的安全。无线局域网用户可以通过使用多种安全手段来防止黑客袭击，比如：通过生物测定学或是一些硬件。一般来说，硬件多指一种小卡片，它可以显示口令，这个口令随时间变化。如果不是本人输入则口令很难正确，当然要进入他人网络也就不可能了。生物测定学的方法是通过对用户指纹和眼睛的鉴定来决定是否允许用户登录。可以使用以上方法来建立一个安全网络，把黑客拒之门外。还可以在需要时锁定电脑，这样就无需担心黑客的骚扰。

IEEE 802.11 标准可以使手提计算机和台式计算机在数百英尺范围内分享内部网络的数据，这种功能是由苹果公司最先推出的。此后。众多计算机生产商都在产品上增加了无线功能。目前，无线局域网在各公司、科研部门、机场，以及其他公共场所的应用越来越广泛。但是，如果 IEEE 802.11 标准的安全漏洞不能及时解决，无线局域网将成为黑客们下一个最主要的攻击目标。

3）未经授权的用户获得存取权

随着 Internet 的不断发展，内部网的用户可以方便地访问外部网，但同时外部网的用户也可以进入到内部网，这就存在着潜在的安全威胁。这种威胁不仅仅存在于 Internet，只要在内部网允许外部用户进入的情况下，就可能产生不安全因素。

比如，远程访问服务器允许在外出差的销售和市场人员通过拨号上网来收取电子邮件，远程办公室通过拨号连接在线站点和 Extranets，它能够将供货商或客户连接到内部网络。所有这些都为黑客、病毒等的入侵提供了可乘之机。事实上，内部网络的最大威胁是来自于网络内部。如果没有正确合适的安全措施，则网络上任何一个已经注册的用户都可以存取数据，现在或以前供职的带有不满情绪的职员已经知道读取甚至修改有价值的公司数据文件的方法，这对数据的安全存在着一定的隐患。因而应当有适当的安全保密产品，对用户的

安全级别正确设置并且随时审查安全程序。

如果已经采取了 VPN 的方法来保护无线客户端，则系统就应当具有很强的认证功能。管理员可以运用 IEEE 802.1x 来防止未经认证的用户登录网络，IEEE 802.1x 也允许管理员选择基于传输层的认证方法，这样可以确保用户只能从经过认证的 AP 进入系统。然而，并不是所有的网络系统都需要强有力的用户认证系统。对于连接提供商来说主要关心的“窃取服务”常常发生在诸如宾馆、机场等热点场合。终究是要由商业模式决定网络访问，而防止未经认证的访问是一种商业需求。

4）无线病毒

目前的 Internet 病毒攻击波也可能会发生在无线装置上，这极有可能是病毒威胁在将来表现出来的新形式。无线病毒能够清除数据、损坏移动电话、PDA 及与无线局域网连接的手提电脑等装置。据报道，世界上出现的第一例无线病毒主要是攻击 Palm 计算公司的 Palm OS 产品，这种病毒是在 2000 年 9 月份首次亮相的。随后，借助短信息服务进行传播的病毒也开始出现，而且会攻击诺基亚公司生产的移动电话及一些 SIM 卡。无线病毒现在还不是十分盛行，主要是因为无线装置的种类繁多。在将来，像 Java 这样跨平台的系统日益增多时，无线病毒的数量及传染速度相应地也会日渐增加。早期的病毒借助电子邮件达到入侵别人的目的。而通过短信接收到的病毒，它会叫用户打开，而一旦打开，用户的手机即出现故障，或者病毒信息被拷贝发送给电话簿中的所有移动电话号码。要清除那些运行像 Palm OS 或 Windows CE 等通用操作系统装置上的病毒并不是很困难，但对移动电话来说则较为费力，因为它们使用的都是一些厂家所独有的软件系统。但同时这一问题能使它本身对于攻击这些无线装置的病毒起到特殊的效果，因为这些装置是人们无法直接与病毒进行接触的应用程序。

2. 针对无线局域网的攻击手段

针对无线局域网的安全缺陷，常见的有以下几个攻击类型。

1）被动攻击——解密业务流

在初始化变量发生碰撞时，一个被动的窃听者可以拦截窃听所有的无线业务流。只要将两个具有相同初始化变量的包进行异或，攻击者就可以得到两条消息明文的异或值，而这个结果可以用来推断这两条消息的具体内容。IP 业务流通常是可以预测的，并且其中包含了许多冗余码，而这些冗余码就可以用来缩小可能的消息内容的范围，对内容的更进一步的推测则可以进一步缩小内容范围，在某些情况下甚至可能确定正确的消息内容。

2）主动攻击——注入业务流

假如一个攻击者知道一条加密消息确切的明文，那他就可以利用这一点来构建正确的加密包。其过程包括：构建一条新的消息、计算 CRC-32、更改初始加密消息的比特数据从而变成新消息的明文、然后将这个包发送到接入点或移动终端，这个包会被当作一个正确的数据包而被接收。这样就将非法的业务流注入网络中，从而增加了网络的负荷。如果非法业务流的数量很大，就会使网络负荷过重，出现严重的拥塞问题甚至导致整个网络完全瘫痪。

3）面向收发两端的主动攻击

在这种情况下，攻击者可以不猜测消息的具体内容而是只猜测包头，尤其是目的 IP 地址，它是最有必要的，这个信息通常很容易获得。有了这些信息，攻击者就可以改变目的 IP

地址,用未经授权的移动终端将数据包发到他所控制的机器上。由于大多数无线设备都与Internet相连,因而这个数据包就会成功的被接入点解密,然后通过网关和路由器向攻击者的机器转发未经加密的数据包,这样就泄露了明文。如果包的TCP头被猜出来的话,那么甚至有可能将包的目的端口号改为80,如果这样的话,它就可以畅通无阻的越过大多数的防火墙。

4) 基于表的攻击

由于初始化向量的数值空间比较小,这样攻击者就可以建一个解密表。一旦知道了某个数据包的明文,它就能够计算出由所使用的初始化变量产生的RC4密钥流。这个密钥流可以将所有使用同一个初始化变量的数据包解密。很可能经过一段时间以后,通过使用上述技术,攻击者能够建立一个初始化变量与密钥流的对照表。这个表只需要很小的存储空间(大约15GB);表一旦建立,攻击者就可以通过无线链路把所有的数据包解密。

5) 广播监听

如果接入点与hub相连而不是与交换机相连,那么任意通过hub的网络业务流将会在整个无线局域网里广播。由于以太网hub向所有与之连接的装置包括无线接入点广播所有数据包,这样,攻击者就可以监听到网络中的敏感数据。

6) DoS攻击

DoS攻击对无线局域网也是一个实际存在的威胁。如果非法业务流覆盖了所有的频段,合法业务流就不能到达用户或接入点。这样,如果有适当的设备和工具的话,攻击者很容易对2.4GHz的频段实施泛洪(flooding)攻击,破坏信号特性,直至导致无线局域网完全停止工作。另外,无绳电话、婴儿监视器和其他工作在2.4GHz频段上的设备都会扰乱使用这个频率的无线局域网。这些拒绝服务可能来自工作区域之外,也可能来自安装在其他工作区域的会使所有信号发生衰减的IEEE 802.11设备。总之,不管是故意的还是偶然的,DoS攻击都会使网络彻底崩溃。

7) “欺诈”AP

无线局域网易于访问且易于配置,这两个特点结合在一起使得网络管理员很头疼。任何用户都可以到计算机商店购买AP并且不需要任何认证而接入公共网络,很多AP在低级管理员那里也设有签字权,一个部门也可以随时脱离某个无线局域网而不需要得到系统确认。

“欺诈”AP可以由终端用户来配置,这一点带来了很大的安全风险。终端用户不是安全专家,并且可能不知道无线局域网带来的安全风险。大部分现存的由War driver所设的配置映射不能使安全特征加入到产品中去,很多AP已经将最小的变化作为默认设置。在这一点上,即使是大公司的终端用户也很难令人相信能做的更好。但是,现在对于这一问题并没有好的解决方法。像NetStumbler这样的工具允许网络管理员在他所在的建筑物中漫游来寻找没有认证的AP,但是在建筑物中漫游来寻找一个新的AP花费大、耗时多。跟踪工具也可以在一个用户与其他企业或机构共存的同一栋建筑物或同一层建筑物中获得其他的AP,其他用户的AP可能部分覆盖该用户所在建筑物层的空间,但是他们的AP不能直接危及该用户网络的安全并且不能引起报替。周期性地巡查网络是确定未经认证的配置的唯一方法,随着网络分析逐渐趋向手工方式,巡查网络时并不需要携带太多的工具。

8）强制服务和执行

无线局域网的传输容量是有限的。建立在 IEEE 802.11b 标准上的无线局域网具有 11Mb/s 的速率，而建立在 IEEE 802.11a 标准上的无线局域网具有 54Mb/s 的速率。这一容量由所有由同一点接入的用户所共享。由于传输媒介层在头上，实际有效的速率大约是正常的一半。不难想象，本地应用可能受到这样容量的限制或者攻击者可以在有限的资源下发动拒绝服务攻击。

对于监视和发现的执行问题，很多 AP 将通过 SNMP 来报告统计数据。但是没有必要详细到使得终端用户抱怨的程度。无线网络分析能够报告某一点的信号质量和网络运行情况，但是无线网络管理员使用的工具才刚刚开始出现。最初的无线分析工具与有线分析工具很相似，像 AirMagnet 的信息处理分析仪这类新产品看来极有可能成为无线网络工程师的必备工具。企业无线局域网管理系统还没有出现，在实际运用中有时能通过连接主干网络的无线局域网靠设置通信流量来扫描到用户的地址。这一点即使不能用来防止拒绝服务攻击也可以有助于防止大量用户垄断某一地区的无线资源。

9）MAC 欺骗和时间段劫持

安全访问一个无线局域网的方法是命令 AP 仅传送来源于已知地址列表中的数据包。MAC 控制地址能够被欺骗，但是在此之前攻击者必须得知一个用户的以太网卡地址。在现实当中，很多无线网卡直接将 MAC 地址标在上面。即使卡的地址可以被保护，但是有效的 MAC 地址列表还是不得不被编辑并保持和分布在每一个 AP。另外，AP 的每一个标记在允许的地址数量上都有一些限制，例如 Lucent 的 Orinoco 拥有最多 490 个 MAC 地址而 Cisco 支持最多大约 2000 个 MAC 地址。

IEEE 802.11 标准并不对帧进行认证。每一帧都有一个源地址，但是无法保证每一帧都确实被发送出去了。比如在传统的以太网中就无法防止伪造帧的源地址，攻击者能够利用欺骗帧来重定向通信流量并且讹用 ARP 表。在许多简单的标准下，攻击者能够观察网络中的某个 MAC 地址并且利用这些地址来发送恶意信息。为了防止这一类攻击，应当在 IEEE 802.11 标准中开发用户认证机制。所有用户都需要认证，未经认证的用户不能访问网络。尽管如此，拒绝服务攻击仍然可能发生，因为无法阻止攻击者访问无线传输层。攻击者能够使用欺骗帧及时间段劫持来发动主动攻击，还能够利用未经认证的 AP。AP 靠广播其信标帧来被鉴别，任何声称是一个 AP 且广播正确服务装置的识别都是网络认证的一部分。然而，攻击者可以容易地假装成一个 AP。这是由于 IEEE 802.11 没有任何功能需要一个 AP 证明它确实是一个 AP。在该点，攻击者能够利用 Man-In-The-Middle（MITM）攻击方法潜入并偷窃信用卡，然后用窃得的合法用户的信用卡来获得网络访问权。

7.2.2 无线局域网安全需求分析

网络的安全性是无线局域网的供货商必须面对的最常见问题之一，作为网络管理员来说关心安全问题是很明智的。但是，心怀不满的雇员、黑客、病毒、工业间谍和其他形式的破坏在网络中并不少见。为了在一个企业或机构或某种环境中使用无线局域网，应当尽量减少目前有关无线局域网产品及标准中存在的内部风险。

保证用户安全需求的切入点是从无线局域网的连接上开始，考虑三个基本的安全服务：审计、认证和保密。

1. 审计

显然,公司或机构网络的最大的威胁来自公司或机构本身。没有正确、适当的安全措施,网络中任何一个已注册的用户都可以存取数据。无论网络管理员在有线网上是否有无线网段,都需要有适当的安全保密产品对用户安全级别正确设置和随时审查安全程序。

网络安全在无线局域网上尤其显得重要,这是因为它很容易在网络内部增加新的 AP。保护无线局域网的第一步就是完成网络审计,实现对内部网络的所有 AP 都做审计,确定欺骗 AP,建立规章制度来约束它们,或者完全从网络上剥离它们。从短期来看,企业应该使用一些能检测出无线局域网网络流量以及无线局域网中的 AP 的网络监控产品或工具,例如像 Sniffer Technologies 和 WildPackets 等厂家的产品。不过,采取的这些措施能达到的安全程度毕竟还是有限的,因为它要求网络管理员要根据无线局域网的信号来检测网络流量,知道网络内部的数据流量情况。企业或机构用户应该形成一个管理制度,保证网络审计成为一个规范化的行为(至少每三个月检测一次),来限制具有欺骗访问行为的站点随意进入无线局域网。

2. 认证

设计一个完善的无线局域网系统,加密和认证是需要考虑的两个必不可少的安全因素。无线局域网中应用加密和认证技术的最根本目的就是使无线业务能达到有线业务同样的安全等级。开放式访问可能产生两个问题,除了未经认证使用的带宽负荷以外,也可能引起法律问题。未经认证的使用者必须服从服务提供商的服务条款,有可能由于一个垃圾邮件而引起 Internet 服务提供商废除连接。无线局域网在应用中存在着较为严重的安全隐患,大多数无线局域网的默认配置允许任意的具有无线网卡的用户在没有任何认证措施的情况下进行访问,这使得访问网络很容易。

无线局域网服务提供商应当十分关注他们的网络配置,如果某个人不经认证而能够轻易地访问无线局域网,则他就能够盗用服务。无论是否经过认证,一旦一个用户获准访问无线局域网,都应当使他只能得到经过认证的服务或其行为得到安全监控。如果这些措施存在缺陷或根本没有这些措施,则一个未经认证的用户就可能轻易地通过无线局域网进入一个个人网络,并且利用网络系统的内部缺陷而完全控制网络。因为基于 WEP 协议的无线局域网安全协议并不是十分可信的,用户必须考虑到提供商可能会留有后门。企业用户也可以配置入侵检测系统(Intrusion Detection System,IDS),作为一种检测欺骗访问站点的前期识别方式。入侵检测系统还能帮助管理员识别特定的、可能存在安全漏洞的访问点或网段,能够帮助网络管理员发现入侵者的物理位置[1]。

用户认证的基础是 2001 年 6 月批准的 IEEE 802.1x 标准。IEEE 802.1x 能够被用来在访问网络前要求用户认证,但是附加的特性需要提供所有的无线局域网要求的密钥管理功能。利用 IEEE 802.1x 标准能够设计出支持相互认证的协议。利用基于 TLS 的方法,在客户提供可信性认证之前 AP 需要提供它的身份证明,并且在空中传输时其可信性由强健的密码技术来保证。在 IEEE 802.11 的 MAC 采取每帧认证之前,时间段劫持攻击将不会得到完全解决。除非时间段劫持攻击是一个连接,就可以在 IEEE 802.11 顶端设计一个密码协议来防止劫持。通过网关控制也可以提供认证需求。网关可以为无线网络建立一个特殊的子网,这些子网拥有在数据包传送前要经过认证的网关而且还可以利用 IEEE 802.1q

标准建立虚拟局域网(Virtual Local Area Network,VLAN)。利用这一标准,选择端口可以通过不同的交换机进入一个子网。只要 VLAN 主干网被交换机所支持,那么即使是地理上分割开来的子网也可以通过交换机互联。即使几个子网作为 VLAN 端口在同一个物理交换机上,使用 VLAN 端口的节点在没有经过路由器或网关的情况下也不能访问其他的子网。一旦 VLAN 被建立起来,一个仅能传输经过认证的用户的通信流量的网关就需要建立起来。由于 VPN 服务器需要认证及提供客户的 IP 地址和密钥,所以在这里也可以利用网关。使用 VPN 服务器作网关不仅需要认证用户,而且要用户独有的密钥加密无线信息流,去除对 WEP 中共享密钥的使用需求。VPN 连接并不理想,理解 VPN 技术、选择 VPN 网关、配置服务器及支持客户等都是不易完成的复杂任务。还有一种特殊的防火墙网关,这种连接仍然使用 VLAN 将聚集的无线信息流量连接到一个网关,但是取代了 VPN,这个网关运行特殊的代码。当一个系统加入该无线局域网时,防火墙/路由器给它一个动态主配置协议。为了进行访问授权,客户应当打开 Web 浏览器。HTTP 协议要求客户经过网关主动发出一个改向的认证页,并且该认证要求发往 Kerberos 服务器。如果认证成功,则将其加入规则文件并在防火墙的 IP 功能表中将其标为“已知”。

从用户的观点来看,他们应该浏览并输入用户身份和口令来获得网络访问权,不需要客户安装和配置。这种方法只能提供认证而不能提供加密,并且对同一时段的用户数量没有限制。这种解决方案在实际应用中是独一无二的,具有很好的效果,它在客户和多家厂商网卡都不用任何改变的情况下允许访问系统。

3. 保密

IEEE 802.11 标准的加密服务使用基于 RC4 算法的 WEP 协议来密封数据帧的有效负载。尽管 WEP 指定了一个 40 位的密钥,但是还有一些厂家使用 104 位的密钥。WEP 的本意并不是提供端对端的加密方案。在 AP 和无线设备上的 WEP 密钥可以被轮换使用,但是由于 IEEE 802.11 标准没有指定密钥管理协议,因此所有的密钥轮换都必须人工进行。同 SSID 一样,轮换 WEP 密钥将影响到所有的 AP 和无线用户,为此将花费网络管理员大量的精力。很多有关 WEP 协议的缺陷已经被指出,该协议仅仅保护与网络的初始连接及用户数据帧,而管理帧和控制帧没有被加密及认证,这就给攻击者利用欺骗帧破坏信息传输留下了可乘之机。早期的 WEP 协议是很容易被一些专门的工具如 AirSnort 及 WEPCrack 等所破解的,但经过改进后已经能防止目前所知的各种攻击了。最新的 WEP 协议作了更进一步的改进,利用密钥管理协议每分钟就更改一次密钥。即使是最繁忙的无线局域网也不会在 15 分钟内为已知的攻击方法产生足够的数据来获得密钥。自从 1999 年 9 月 IEEE 批准了 802.11b 标准以来,WEP 协议就成为无线局域网上应用的主要的加密机制,用来对无线局域网上的数据流进行加密。不过目前许多企业并没有启动 WEP,主要是因为 WEP 的密钥管理和配置起来过于繁琐。尽管 META Group 已经承认了一些与 WEP 有关的漏洞,不过最近报道的一些攻击行为证明,该保密机制的漏洞要比想象中的还要多。所以企业用户必须依据使用环境的机密要求程度,来对使用的应用软件进行评估。IEEE 802.11b 标准在客户端和 AP 之间靠 WEP 提供保密通信。在 WEP 下面,已经给出 AP 的所有用户共享同样的加密密钥。为了在一个范围内具有可移动性,所有的 AP 应当被设置成使用同样的密钥并且所有的客户也具有同样的加密密钥。另外,数据头保持未加密状态,以便任何人都能够看到数据传输的源和目的。

7.3 安全技术

安全技术包括两个方面：访问控制和保密性。访问控制确保敏感的数据仅由获得授权的用户访问，保密性则确保传送的数据只被目标接收人接收和处理。由于无线局域网通过电磁波在空中传播数据，所以在接入点覆盖区域内的几乎任何一个无线局域网用户都能接触到这些数据。无论接触数据者是在另外一个房间、另外一层楼或是在本建筑物之外，要将无线局域网发射的数据只传送给一个目标接收者是根本不可能实现的，无线就意味着会让人接触到数据。无线局域网用户需要更强大的安全解决方案，来解决数据传输过程中的安全性。因此，数据保密性就成了无线局域网应用中一个需要非常关注的方面。

由于无线局域网采用公共的电磁波作为载体，因此与有线线缆不同，任何人都有条件窃听或干扰信息，因此在无线局域网中，网络安全很重要。下面介绍无线局域网采用的一些安全技术。

7.3.1 服务装置标识符

无线客户端必须出示正确的 SSID 才能访问无线接入点 AP，利用服务装置标识符(Service Set Identifier，SSID)可以很好地进行用户群体分组，避免任意漫游带来的安全和访问性能的问题。因此可以认为 SSID 是一个简单的口令，从而为无线局域网提供一定的安全性。然而无线接入点 AP 向外广播其 SSID，使安全程度下降。另外，一般情况下，由用户自己配置客户端系统，所以很多人都知道该 SSID，非法用户也可能得到 SSID。况且有的厂家支持任何 SSID 方式，只要无线客户端处在 AP 范围内，那么它都会自动连接到 AP，这将绕过 SSID 的安全功能。

7.3.2 物理地址过滤

物理地址过滤：每个无线客户端网卡都由唯一的物理地址标识，因此可以在 AP 中手工维护一组允许访问的 MAC 地址列表，实现物理地址过滤。物理地址过滤属于硬件认证，而不是用户认证。这种方式要求 AP 中的 MAC 地址列表必须随时更新，目前都是手工操作；如果用户增加，则扩展能力很差，只适合于小型网络规模。另外，非法用户利用网络侦听手段很容易窃取合法的 MAC 地址，而 MAC 地址并不难修改，因此非法用户完全可以盗用合法的 MAC 地址进行非法接入。

7.3.3 直接序列扩频技术

在 IEEE 802.11b 高速标准中，大多数的厂商都使用直接序列扩频技术(DSSS)作为物理层的选择。DSSS 将每一个位信息传送之后再附加另外一个位，称为 Chip，提供容错的功能以及信息传递的一致性，Chip 也让信息的传输更加安全。尽管如此，黑客还是可以使用扩频分析仪去截取无线电波，也可以用特定的无线网卡去搜寻各频道内的数据，进而加以分析与破解。为了克服这个问题，就要将无线传输中的信息加密，这样即使信息被黑客中途拦截也无法破解。在 DSSS 方式中，信号可以跨越很宽的频段，数据基带信号的频谱被扩展几

倍到几十倍再被搬移至射频发射出去。

这一做法虽然牺牲了频带带宽，但由于其功率密度随频谱的展宽而降低，甚至可以将通信信号淹没在自然背景噪声中，因此其保密性很强。要截获或窃听、侦察这样的信号是非常困难的，除非采用与发送端相同的扩频码与之同步后再进行相关的检测，否则对扩频信号是无能为力的。

7.3.4　扩展服务集标识符

在每一个 AP 内都会写入一个服务区域认证 ID，每当端点要连上 AP 时，AP 会检查其扩展服务集标识符（ESSID）是否与其相同，如果不符就拒绝给予服务。譬如一个 AP 上的 ESSID 是“MY-Wireless-Net”，而想连接的使用者却不知道 AP 的 ESSID，则其就会被拒绝。

7.3.5　开放系统认证

开放系统认证是 IEEE 802.11 标准的一种默认认证协议，它对任何要求授权的客户进行认证。正如其名，无论谁请求认证都会被对方通过，实质上，它是一个空认证过程。试验表明，在进行网络连接时，终端之间确实采用这种方法进行相互认证，而且即使采用了 WEP 协议进行认证，这种认证的管理帧也是可以随意地在网络中传输，并且不受任何阻碍。

7.3.6　共享密钥认证

共享密钥认证使用一个标准的询问和响应帧格式，其中包含一个用于认证的共享密钥。请求认证的终端发送一个认证请求管理帧，表明它请求进行共享密钥认证。认证请求的接收端则发送一个包含 128 个字节询问正文的认证管理帧给发送端作为响应。这个询问正文由 WEP 的伪随机序列发生器产生，其中包括共享密钥和一个随机初始化向量（IV）。一旦发送端收到管理帧，它将询问正文的内容复制到一个新的管理帧的正文中，然后 WEP 协议使用共享密钥和新的 IV 来加密这个新的管理帧，最后将加密后的管理帧发送到接收端。接收端将收到的帧解密，首先确定 32 比特的 CRC 完整校验值是否正确，然后再确定询问正文是否与第一条消息相同。如果全部正确，这样认证就成功了。如果认证成功，发送端和接收端交换一下角色，再进行一次上述的过程，以确保双方相互认证。共享密钥认证利用一个标准请求和应答一起共享秘密钥来提供认证。当某个站点要获取授权时，它发出一个认证请求管理帧表明它想使用共享密钥认证，收到请求后，应答者将包含着 128 个 8 位数的应答文本的认证帧发回给请求者，应答文本由用 WEP 伪随机数产生器利用共享密钥和一个随机初始化向量生成。一旦请求者从应答者那里收到管理帧，它就将管理帧的内容拷贝生成一个新的管理帧。新的管理帧随后被 WEP 用共享密钥和请求者新选的初始化向量加密，然后将加密后的管理帧发送给应答者。应答者对收到的帧解密并校验，然后将其与第一次发出信息的请求文本进行比较，如果相同则认证成功。如果认证成功，则请求者与应答者互换角色并重复这一过程以加强相互认证。如果认证成功，则状态码置零；如果认证不成功则返回一个错误值。要素标识符中包含着请求文本，长度域标识出请求文本的长度，最长是 128。请求文本包含了随机请求串。

7.3.7 封闭网络访问控制

朗讯(Lucent)公司定义了一个称为封闭网络的特殊的访问控制机制。在该机制中,网络管理员可以使用开放的或封闭的网络。在开放的网络中,任何人都允许连接访问网络;而在封闭的网络中,只有那些知道网络名称或 SSID 的用户才可连接。在这里,网络名实际上就作为一个共享密钥。但是,这个机制只适合于朗讯自己的产品。

7.3.8 访问控制列表

在软件开发上采用的另一个保证安全的机制是基于用户以太网 MAC 地址的访问控制机制,但是这个机制并没有在标准中定义。可以将无线局域网只设定为给特定的节点使用,因为每一张无线网卡都有一个唯一的 MAC 地址,只要将其分别输入 AP 即可。相反的,如果有网卡被偷或发觉有存取行为异样,也可以将这些 MAC 地址输入,禁止其再次使用。利用这个存取控制机制,如果有外来的不速之客得知公司使用的无线局域网 ESSID 也一样会被拒绝在外。每一个 AP 都可以用所列出的 MAC 地址来限制网络中的用户数,如果用户的 MAC 地址存在于列表中,那么就允许它访问网络;如果不在列表中,就不允许它访问。

7.3.9 密钥管理

就 IEEE 802.11 标准而论,其实并没有实现严格意义上的密钥管理,只有少数开发商在他们的高端产品中实现了某一形式的密钥管理或密钥协议,所有开发商都没有提供足够的信息来确定产品所保证的安全等级。

IEEE 802.11 标准提出两种使用 WEP 密钥的方法。第一种方法提供了一个 4 个密钥的窗口,终端或 AP 能够使用任何一种密钥来解密数据包。然而,在传输数据时,只能手动输入 1 个密钥——默认密钥。第二种方法叫做密钥映射表,这个方法规定每个唯一的 MAC 地址都有 1 个单独的密钥。根据 IEEE 802.11 标准,密钥映射表的大小至少包含 10 个条目,最大的容量则取决于芯片的设置。每个用户使用各自单独的密钥可以减少密码攻击的可能性,但是实施一个合理的密钥周期仍然是一个问题,因为密钥只能手动改变。

密钥管理是 IEEE 802.11 的一个失误的地方,它好像是留给厂商的作业题,但事实上只有少数几个厂商在他们的高端产品中设置了密钥管理或密钥协商的功能。实际上,没有一个厂商能提供充分的信息以确定他们的产品所能提供的安全保证的水平。在有些情况下,由于使用众所周知的脆弱的协议如没有认证的 Diffie-Hellman 密钥交换协议等,厂商的"解决方案"反而使得安全性变得更糟[2]。

7.3.10 虚拟专用网

虚拟专用网(Virtual Private Networks,VPN)技术是目前快速增长的一种安全技术,用来在公共网络基础设施(public network infrastructures)上建立一个虚拟的专用通道,进行安全的数据传输。近年来,VPN 技术已经使得企业可以利用 Internet 进行远程访问(remote access)。目前,VPN 技术主要应用在下面三种不同的场合:远程访问,局域网间

(LAN-to-LAN)的连通和特殊网络。对于较大规模和安全等级高的商业网络来说,VPN 是替代 WEP 和 MAC 地址过滤的较为理想的无线接入的安全方案。在 VPN 安全方案中,VPN 为接入用户提供一条专用的安全接入隧道到内部网络。常用的隧道有 PPT、L2PP,网络结构为标准的集中认证结构,如 Radius 服务器认证。客户端与内部网络被 AP 和 VPN 服务器之间的局域网和 VPN 服务器隔开。VPN 服务器负责客户端的认证和传输加密,同时作为内部网络的网关。

VPN 方案的优点有:

- 适用于用户众多的大规模网络。
- 对 AP 和客户端的管理需求小,而 VPN 服务器可集中管理。
- 客户端与内部网络隔离,通信前必须经过 VPN 认证。
- WEP 密钥和 MAC 地址列表的管理成为可选项。
- 统一的用户界面。

VPN 使用起来也有一些缺点:

- 现阶段 WLAN 的 VPN 安全方案不支持广播功能。大规模网络通过广播功能从信息源向多用户传送视频、音频等信息,如果通过多个点对点传送来实现,将占用网络的很多带宽,而广播能更有效地利用网络带宽在骨干线路上传输这些信息。
- 当移动终端从一个 VPN 服务器所在的子网漫游到另一个子网,用户需重新登录。同样,当客户端从待机模式重新激活时也需要重新登录。

7.3.11 RADIUS 服务

RADIUS(Remote Authentication Dial-In User Service,远程拨号接入用户认证协议)[4]以客户机/服务器模式工作,实现了对访问网络用户的身份认证、授权、计费等增强的服务功能。其客户端多为网络访问服务器(NAS),主要用于将用户信息传递给 RADIUS 服务器,RADIUS 服务器对用户进行认证,并返回用户的网络访问配置信息。RADIUS 协议特点:

1. 客户机/服务器模式

RADIUS 协议的客户端通常是网络接入服务器(NAS)。客户端的任务是把用户的信息(如账号、口令等)传给指定的 RADIUS 服务器,并接受服务器的响应。RADIUS 服务器的任务主要是:接受用户的连接请求;对用户身份进行认证;返回用户接入网络的所有配置信息。同时,RADIUS 服务器还可以作为其他 RADIUS 服务器或异种认证服务器的代理客户。

2. 网络安全方面

客户端和 RADIUS 服务器之间的交互经过了共享密钥法认证。另外,为防止他人经过传输线路获得用户口令,在传输过程中对口令进行了加密。

3. 灵活的认证机制

RADIUS 服务器支持多种认证方法。当用户提供用户名和原始口令时,它可以支持 PPP、CHAP、UNIX Login 和其他的认证机制。

4. 协议的可扩展性

所有的交互都包含可变长的属性字段。为满足实际需要，用户可以加入新的属性值而不会对原协议有任何影响。

大公司的远程用户常常通过 RADIUS 实现网络认证登录。企业的 IT 网络管理员能够将无线局域网集成到已经存在的 RADIUS 架构内来简化对用户的管理。这样不仅能实现无线网络的认证，而且还能保证无线用户与远程用户使用同样的认证方法和账号。

7.3.12 入侵检测系统

入侵检测系统(Intrusion Detection System，IDS)是一种用来检测是否存在未经授权的用户试图访问网络或访问已经访问网络，甚至已经危及网络安全的有效工具。对于 WLAN 而言，IDS 可能基于主机(host-based IDS)，也可能基于网络(network-based IDS)。

基于主机的 IDS 为特殊的漏洞或重要的系统增加了一个安全层。基于主机的代理安装于单独的系统中(例如，数据库服务器)，用于监视审计记录及可疑行为的系统日志(例如，多次错误登录或更改文件的使用权限等)。基于主机的代理也可以使用校验和(checksum)定期查看系统文件的变动情况。虽然基于主机的代理主要功能在于生成日志，分析事件和在紧急情况下发警报等，但是在一些情况下，代理也可以中断对系统的攻击。

基于网络的 IDS 用来实时(或尽可能实时)地按数据包监视 LAN(或部分 LAN)上的网络通信，确定传输的数据是否与预定的攻击特征(与已知攻击特征相匹配的行为)相一致。例如，TearDrop DoS 拒绝服务攻击会以类似于碰撞目标系统的方法来发送数据包片断。网络监听可以识别出与这种攻击特征一致的数据包，并采取诸如切断网络会话等保护措施，同时向管理员发 E-mail 报警或采取其他指定的行动。对于像 SSL(Security Socket Layer)网络会话或 VPN 连接等所涉及的加密通信，基于主机的 IDS 系统要优于基于网络的 IDS。这是因为代理位于组件本身，所以基于主机的 IDS 系统能够检查加密后的数据。相比而言，基于网络的 IDS 系统不能解密数据，因此加密的数据会不经检测进行传输。

无线入侵检测系统用于集中式和分散式两种。集中式无线入侵检测系统通常用于连接单独的传感器，搜集数据并转发到存储和处理数据的中央系统中。分散式无线入侵检测系统通常包括多种设备来完成 IDS 的处理和报告功能。分散式无线入侵检测系统比较适合较小规模的无线局域网，因为它价格便宜和易于管理。多线程的处理和报告的传感器管理比集中式无线入侵检测系统花费更多的时间。

无线局域网通常被配置在一个相对大的场所。像这种情况，为了更好地接收信号，需要配置多个无线基站(Wireless Access Points，WAPs)，在无线基站的位置上部署传感器，这样会提高信号的覆盖范围。由于这种物理架构，大多数的黑客行为将被检测到。另外的好处就是加强了同无线基站的距离，从而，能更好地定位黑客的详细地理位置。

7.3.13 个人防火墙

由于公共无线网络上的资料通常没有如内部资料那样的保护措施，所以很容易遭到攻击。个人防火墙针对某些攻击提供了一些保护措施。个人防火墙是基于软件的解决方案，安装在客户端，由客户端管理或进行集中式管理。对于由客户端管理的防火墙，个人用户无

需遵循特定的规则即可方便地进行配置，所以最适合低端用户。而集中式管理防火墙可以使 IT 部门进行配置或远程管理，提供了更高水平的安全防护。集中式管理的解决方案允许企业根据已知漏洞来修改客户端防火墙，并为所有的远程用户提供一致的安全策略。一些这类高端产品也拥有 VPN 和安全审计的功能。虽然个人防火墙可以提供一些保护措施，但是它们也不能够防范高级的攻击方法。

7.3.14　基于生物特征识别

基于生物特征识别(biometrics)包括指纹/掌纹扫描器、视网膜和虹膜扫描器、面部扫描器和语音图谱扫描器等。单独使用基于生物特征识别技术或与其他安全方案一起使用可以提供另外一层保护措施。例如，安全性要求很高的企业可以将生物特征识别系统集成到无线智能卡，带有无线网卡的笔记本电脑和其他一些无线设备中，在过去常使用用户名/密码来访问无线网络的地方使用基于生物特征识别系统。此外，生物特征识别技术可以和 VPN 结合起来提供身份认证和数据保密性。

7.3.15　双因素身份认证

一种可选的身份认证方案是双因素认证。双因素认证(two-factor authentication)的一种形式是使用对称密码算法每分钟生成一个新码字，这个码字和用户个人识别号(Personal Identification Number，PIN)搭配使用，且只能使用一次。双因素认证的另一个形式是将用户智能卡和个人识别号搭配使用。因此，双因素认证需要有智能卡阅读机或个人识别号认证服务器。

7.3.16　智能卡

智能卡(smart card)可以为 WLAN 增加另外一层保护。可以使用和用户名、密码相关的智能卡，也可以使用双因素身份认证的智能卡。例如，将智能卡与基于生物特征识别结合起来使用。在无线网络中，智能卡提供了另外一种身份认证形式。在要求除简单的用户名和密码之外的身份认证时，智能卡非常有效。用户的数字证书和其他信息都存在智能卡上，通常只要求用户记住个人识别号即可。由于智能卡便于携带，所以用户可以从不同的地方安全访问无线网络。如同身份认证软件解决方案一样，智能卡也被集成到 WLAN 系统，用以增强整个系统的安全性。此外，用户应该充分认识智能卡所提供的安全性，单独使用智能卡不可能解决 802.11 安全中的所有问题。

7.4　安全协议

目前，WLAN 业务的需求日益增长，但是相应的安全措施却无法令人满意。最初人们在研究无线网络的安全问题时，理所当然地把原来应用于有线网络的安全协议植入到无线网络中去，但是这种移植的效果，从 WLAN 安全标准的发展情况看，还远远未达到要求。IEEE 正致力于消除 WLAN 的安全问题，因此发展了一系列的安全协议。

7.4.1 WEP 协议

为了保障无限局域网中实体间的通信遭受窃听和其他攻击，802.11 协议中定义了WEP 子协议，它规定了对无线通信数据进行加密的方法，并对无线网络的访问控制等方面做出了具体的规定。WEP 设计的思想是：通过使用 RC4 流密码算法加密来保护数据的机密性；通过移动站 Station 与访问点 AP 共享同一密钥实施接入控制；通过 CRC-32 循环冗余校验值来保护数据的完整性。

1. RC4 加密算法

RC4 是 Ron Rivest 在 1987 年为 RSA 数据安全公司开发的，它是一种可变密钥长度的流密码，该算法以 OFB 方式工作，密钥序列与明文相互独立。它有一个 8×8 的 S 盒：S_0，S_1，…，S_{255}。所有项都是数字 0～255 的置换，并且这个置换是一个可变长度密钥的函数。它有两个计数器：i 和 j，初值均为 0。要产生一个随机字节，需要经过以下计算：

$$i = (i + 1) \bmod 256$$

$$j = (j + S_i) \bmod 256$$

交换 S_i 和 S_j：

$$t = (S_i + S_j) \bmod 256$$

$$K = S_t$$

字节 K 与明异或得到密文，或与密文异或得到明文，加密速度是 DES 的 10 倍。S 盒的初始化过程如下：首先将其进行线性填充 $S_0=1, S_1=1, \cdots, S_{255}=255$。然后用密钥填充另一个 256 字节的数组，密钥不够长时可重复利用给定密钥以填满整个数组：$K_0, K_1, \cdots, K_{255}$。将计数器 j 设为 0，执行下述程序：

$$\text{for } i = 0 \text{ to } 255$$

$$j = (j + S_i + K_i) \bmod 256$$

$$\text{swap}(S_i, S_j)$$

以上就是全部的描述，RSA 数据安全公司宣称该算法对差分攻击和线性分析是免疫的，没有短循环，并且具有高度非线性，目前尚无公开的分析结果。它大约有 $256! \times 256^2 = 2^{1700}$ 种可能的状态，S 盒在使用中慢慢改变：i 保证每个元素的改变，j 保证元素随机的改变，算法简单，易于编程实现。可以设想利用更大的 S 盒和更长的字，若要用 16×16 的S 盒，初始化过程将会很漫长。RC4 流密码算法是无线局域网的安全协议 WEP 和 TKIP 中采用的加密算法，理解它的基本工作原理有助于我们对这些安全协议进行分析。

2. WEP 协议的基本原理

1）WEP 数据的加密过程

首先计算原始数据包中明文数据的 CRC-32 冗余校验码。设消息为 M，CRC 校验和为 ICV，则得到传输明文数据为 $P=\langle M, \text{ICV}\rangle$。接下来是选用 RC4 算法进行加密。将 24bit 的初始化矢量 IV 与共享密钥 Key 连接起来构成种子密钥，采用 RC4 算法生成密钥序列 RC4(IV，Key)，再将密钥序列与传输明文序列进行异或得到密文即相应的密文为 $C=P \oplus \text{RC4}(\text{IV}, \text{Key})$。发送方将 IV 以明文形式和密文 C 一起发送。

2）WEP 数据的解密过程

在密文 P 传送到接收方以后，接收方从数据包中提取出 IV 和密文，将 IV 和持有的密钥 Key 一起送入采用 RC4 算法的伪随机数发生器得到解密密钥流，该解密密钥流实际上与加密密钥流相同，再将解密密钥流与密文相异或，就得到了原始明文 M 和它的 CRC 校验和 ICV。解密过程可以表示为下式：

$$\{M,\mathrm{ICV}\}=C\oplus \mathrm{RC4}(\mathrm{IV},\mathrm{Key})=\{M,\mathrm{ICV}\}\oplus \mathrm{RC4}(\mathrm{IV},\mathrm{Key})\oplus \mathrm{RC4}(\mathrm{IV},\mathrm{Key})$$

3）WEP 数据的完整性保护

为了防止数据在无线传输过程中遭到篡改，WEP 采用 CRC-32 循环冗余校验和来保护数据的完整性。发送方在发出数据包前要计算明文的 CRC-32 校验和 ICV，并将明文 P 与 ICV 一起加密后发送。接收方收到加密数据后先对数据进行解密，然后计算解密出的明文的 CRC-32 校验和，并将计算值与解密出的 ICV 进行比较，若二者相同，则认为数据在传输过程中没有被篡改，否则认为数据已被篡改过，丢弃该数据包。

4）WEP 规定的访问控制

WEP 协议规定了两种认证方式：开放系统认证和共享密钥认证。开放系统认证的实质是不进行用户认证，任何接入 WLAN 的请求都被允许。共享密钥认证是通过检验 AP 和 Station 是否共享同一密钥来实现的，该密钥就是 WEP 的加密密钥。此认证采用 Challenge-Response 方式，当移动站 Station 想要接入无线网络时，它搜索距离最近的访问点 AP。找到访问点 AP 以后，移动站 Station 向访问点 AP 发送一个接入请求，访问点 AP 接收到 Station 的请求以后，向 Station 发送一个随机数。Station 用双方的共享密钥和上述的加密方法对收到的随机数加密，将密文回送给访问点 AP。AP 再用双方的共享密钥对密文进行解密，将解密结果与发送的随机数相比较，若相同则验证了 Station 是合法用户，允许其接入；否则，拒绝该 Station 的接入请求。

3. WEP 协议的安全性

WEP 虽然通过加密提供网络的安全性，但存在许多缺陷。美国加州大学伯克利分校的 Borisov，Goldberg and Wagner 最早发表论文[5] WEP 协议中存在的设计失误，接下来信息安全界的研究人员发表了大量论文详细讨论了 WEP 协议中的安全缺陷，并与工程技术人员协作在实验中破译了使用 WEP 协议加密的无线传输数据。WEP 安全缺陷具体体现在以下几个方面[5,6]。

1）缺少密钥管理

用户的加密密钥必须与 AP 的密钥相同，并且一个服务区内的所有用户都共享同一把密钥。WEP 标准中并没有规定共享密钥的管理方案，通常是手工进行配置与维护。由于同时更换密钥费时与困难，所以密钥通常长时间使用而很少更换。倘若一个用户丢失密钥，则将殃及到整个网络。

2）ICV 算法不合适

WEP ICV 是一种基于 CRC-32 的用于检测传输噪音和普通错误的算法。CRC-32 是信息的线性函数，这意味着攻击者可以篡改加密信息，并很容易地修改 ICV，使信息表面上看起来是可信的。能够篡改即加密数据包使各种各样的非常简单的攻击成为可能。

3）RC4 算法存在弱点

在 RC4 中，人们发现了弱密钥。所谓弱密钥，就是密钥与输出之间存在超出一个好密

码所应具有的相关性。在24位的IV值中,有9000多个弱密钥。攻击者收集到足够的使用弱密钥的包后,就可以对它们进行分析,只需尝试很少的密钥就可以接入到网络中。目前能够截获WLAN中无线传输数据的硬件设备已经能够在市场上买到,可以对截获数据进行解密的黑客软件也已能够在Internet上下载,如airsort[7]、wepcrack[8]等,WEP协议面临着前所未有的严峻形势,对它进行修订已经是迫在眉睫了。

7.4.2 WEP的改进方案TKIP

TKIP基于RC4,在WEP的基础上增加了一些新的算法,主要有:

(1) MIC码(Message Integrity Code)。它的作用在于验证消息的真伪,如果在一秒之内发现了与MIC码不匹配的数据帧,TKIP认为受到了攻击。客户端将删除密钥,取消与AP的关联,等待一分钟后再与AP重新关联。

(2) 新的IV起始算法。为防止"重放"攻击,TKIP建立了MIC密钥与分组序列号的关联。一旦MIC密钥更换,分组序列号就重新起始。

(3) 密钥混合。生成临时密钥,取代WEP中使用的共享密钥。临时密钥有一个的生命周期,过期后将被新的临时密钥取代。

(4) 三级密钥层次下的密钥更新。TKIP密钥混合最多能生成216个IV,因此每发送216个分组,TKIP就需要更新一次临时密钥。

目前Wi-Fi推荐的无线局域网安全解决方案WPA(Wi-Fi Protected Access)以及制定中的IEEE 802.11i标准均采用临时密钥完整性协议TKIP(Temporal Key Integrity Protocol)作为一种过渡安全解决方案。然而WEP算法的安全漏洞是由于WEP机制本身引起的,与密钥的长度无关,即使增加加密密钥的长度,也不可能增强其安全程序,初始化向量IV长度的增加也只能在有限程度上提高破解难度,比如延长破解收集时间,并不能从根本上解决问题,因为作为安全关键的加密部分,TKIP没有脱离WEP核心机制。甚至TKIP更易受攻击,因为它采用了Kerberos密码,常常可以用简单的猜测方法攻破[9]。另一个严重问题是加/解密处理效率问题没有得到任何改进,甚至更差。Wi-Fi联盟和IEEE 802委员会也承认,TKIP只能作为一种临时的过渡方案,而不是最终方案。

7.4.3 认证端口访问控制技术(IEEE 802.1x)

IEEE 802.1x协议,称为基于端口的访问控制协议(Port Based Network Access Control Protocol),是由IEEE 2001年6月提出来的符合IEEE 802协议集的局域网接入控制协议。主要是为了解决无线局域网用户的接入认证问题,能够在利用IEEE 802局域网优势的基础上提供一种对连接到局域网用户的认证。802.1x的核心是可扩展认证协议EAP(Extensible Authentication Protocol)[3],实质是对以太网端口进行鉴权,可应用于无线和有线以太网络。

802.1x标准是需要网络服务的系统和网络之间的认证对话,这个对话采用了IETF的EAP协议。802.1x标准由申请者(suppliant)和认证端(authenticator)的端口接入实体PAE(Port Access Entity)、EAP封装协议(EAPOL或EAPOW)和远程用户接入认证服务器(Radius Access Server)组成。802.1x标准对我们传统概念的网络端口再予以定义,并对

它增加了认证功能。在 802.1x 标准中最主要的组件是网络接入端口(network access port),它可以是物理网络接口或 MAC 地址,在端口的上一层是端口接入实体 PAE;另一个组件是逻辑控制,管理哪些数据包允许通过,发送给其他设备,哪些数据包被拒绝。

802.1x 标准中申请者(suppliant)和认证端(authenticator)都支持双端口和 PAE 的概念。在基础结构无线局域网络中,无线工作站 STA 就是申请者,接入点 AP 就是认证端。在 802.1x 中认证端即接入点有两种类型的端口(port):可控制端口(controlled port)和不可控制端口(uncontrolled port)。可控制端口只有在无线工作站被认证后,认为是合法用户时才打开,通过该端口对合法用户提供网络服务;不可控制端口一直处于打开状态,对所有的网络流量进行过滤只让建立认证的数据包通过,即 EAPOL 数据包。这种双端口模式对不支持 802.1x 标准用户具有兼容性;通过设置一个管理项可允许它们的流量通过不可控制端口。

802.1x 标准的另一个重要组成部分是认证服务器 AS。认证端(authenticator)充当申请者和认证服务器 AS 之间的 EAP 代理,也就是说,认证端接收来自申请者的 EAPOL 数据包,然后把 EAP 数据包通过例如 RADIUS 等高层协议转发给 AS,同时它把所有来自 AS 的 EAP 数据包通过 EAPOL 协议转发给申请者。通过这种方法 AS 对申请者进行认证,认证通过后产生一个共享会话密钥,同时并把认证结果和会话密钥发送到认证端。

7.4.4　IEEE 802.11i

为了解决 WLAN 技术自身存在的安全上的缺陷,给无线用户提供足够的安全保护,IEEE 802.11 的 i 工作组致力于制订被称为 IEEE 802.11i 的新一代安全标准,这种安全标准为了增强 WLAN 的数据加密和认证性能,定义了 RSN(Robust Security Network)的概念,并且针对 WEP 加密机制的各种缺陷做了多方面的改进。

802.11i 草案的目标是实现数据机密性、身份识别、接入控制、抗重放攻击和数据完整性校验。其中,机密性和数据完整性校验以及抗重放攻击由 TKIP 算法或者基于 AES 的算法一次性实现。

IEEE 802.11i 规定使用 802.1x 认证和密钥管理方式,在数据加密方面,定义了 TKIP(Temporal Key Integrity Protocol),CCMP(Counter-Mode/CBC-MAC Protocol)和 WRAP(Wireless Robust Authenticated Protocol)三种加密机制。其中 TKIP 采用 WEP 机制里的 RC4 作为核心加密算法,可以通过在现有的设备上升级固件和驱动程序的方法达到提高 WLAN 安全的目的。CCMP 机制基于 AES(Advanced Encryption Standard)加密算法和 CCM(Counter-Mode/CBC-MAC)认证方式,是实现 RSN 的强制性要求。由于 AES 对硬件要求比较高,因此 CCMP 无法通过在现有设备的基础上进行升级实现。WRAP 机制基于 AES 加密算法和 OCB(Offset Codebook),是一种可选的加密机制。

到目前为止,802.11i 中可以认为加密、完整性校验和抗重放攻击等已经基本固定。但是在身份识别和密钥管理这一部分,以及和其他协议(802.11e、802.1x)的融合这一部分还会发展。

7.4.5　WPA

WPA(Wi-Fi Protected Access)推出之前,802.11 标准存在的安全缺陷已被广泛关注,

而正在制定中的 802.11i 要到 2003 年底和 2004 年初之间才能完成最终的标准化工作。市场对于提高 WLAN 安全的需求是十分紧迫的，IEEE 802.111 的进展并不能满足这一需要。在这种情况下，Wi-Fi 联盟制定了 WPA(Wi-Fi Protected Access)标准。这一标准采用了 IEEE 802.11i 的草案，保证了与未来出现的协议的前向兼容。

WPA 采用了 802.1x 和 TKIP 来实现 WLAN 的访问控制、密钥管理与数据加密。WPA 系统在工作的时候，先由 AP 向外公布自身对 WPA 的支持，在 Beacons/Probe Response 等报文中使用新定义的 WPA 信息元素(Information Element)，这些信息元素中包含了 AP 的安全配置信息(包括加密算法和安全配置等信息)。STA 根据收到的信息选择相应的安全配置，并将所选择的安全配置表示在其发出的 Association Request 和 Re-Association Request 报文中。WPA 通过这种方式来实现 STA 与 AP 之间的加密算法以及密钥管理方式的协商。

支持 WPA 的 AP 工作需要在开放系统认证方式下，STA 以 WPA 模式与 AP 建立关联之后，如果网络中有 RADIUS 服务器作为认证服务器，那么 STA 就使用 802.1x 方式进行认证；如果网络中没有 RADIUS，STA 与 AP 就会采用预共享密钥(Pre-Shared Key，PSK)的方式。STA 通过了 802.1x 身份验证之后，AP 会得到一个与 STA 相同的 Session Key，AP 与 STA 将该 Session Key 作为 PMK(Pairwise Master Key，对于使用预共享密钥的方式来说，PSK 就是 PMK)。随后 AP 与 STA 通过 EAPOIrKEY 进行 WPA 的四次握手(4-Way Handshake)过程，如图 7-6 所示。

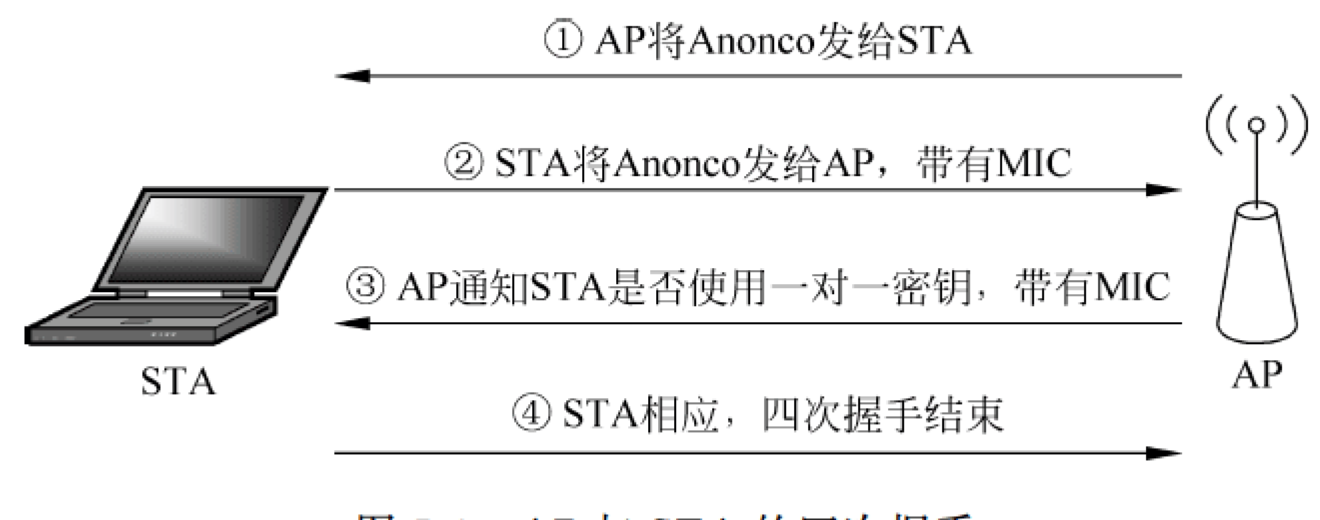


图 7-6　AP 与 STA 的四次握手

在这个过程中，AP 和 STA 均确认了对方是否持有与自己一致的 PMK，如果不一致，四次握手过程就告失败。为了保证传输的完整性，在握手过程中使用了名为 MIC(Message Integrity Code)的检验码。在四次握手的过程中，AP 与 STA 经过协商计算出一个 512 位的 PTK(Pairwise Transient Key)，并将该 PTK 分解成为五种不同用途的密钥，如图 7-7 所示。

其中前 128 位用做计算和检验 EAPOLKEY 报文的 MIC 的密钥，随后的 128 位作为加密 EAPOL-KEY 的密钥；接下来的 128 位作为 AP 与该 STA 之间通信的加密密钥的基础密钥(即由该密钥再经过一定的计算后得出的密钥作为二者之间的密钥)；最后两个 64 位的密钥分别作为 AP 与该 STA 之间报文的 MIC 计算和检验密钥。

由 PTK 分解出来的这一组(五个)密钥是 AP 与该 STA 之间使用的密钥(所以也叫每用户密钥，用于 AP 与 STA 之间的单播报文的加密)。在确认双方所持的 PMK 一致后，AP 会根据自身是否支持每用户密钥的能力来指示 STA 是否安装并使用这个每用户密钥。

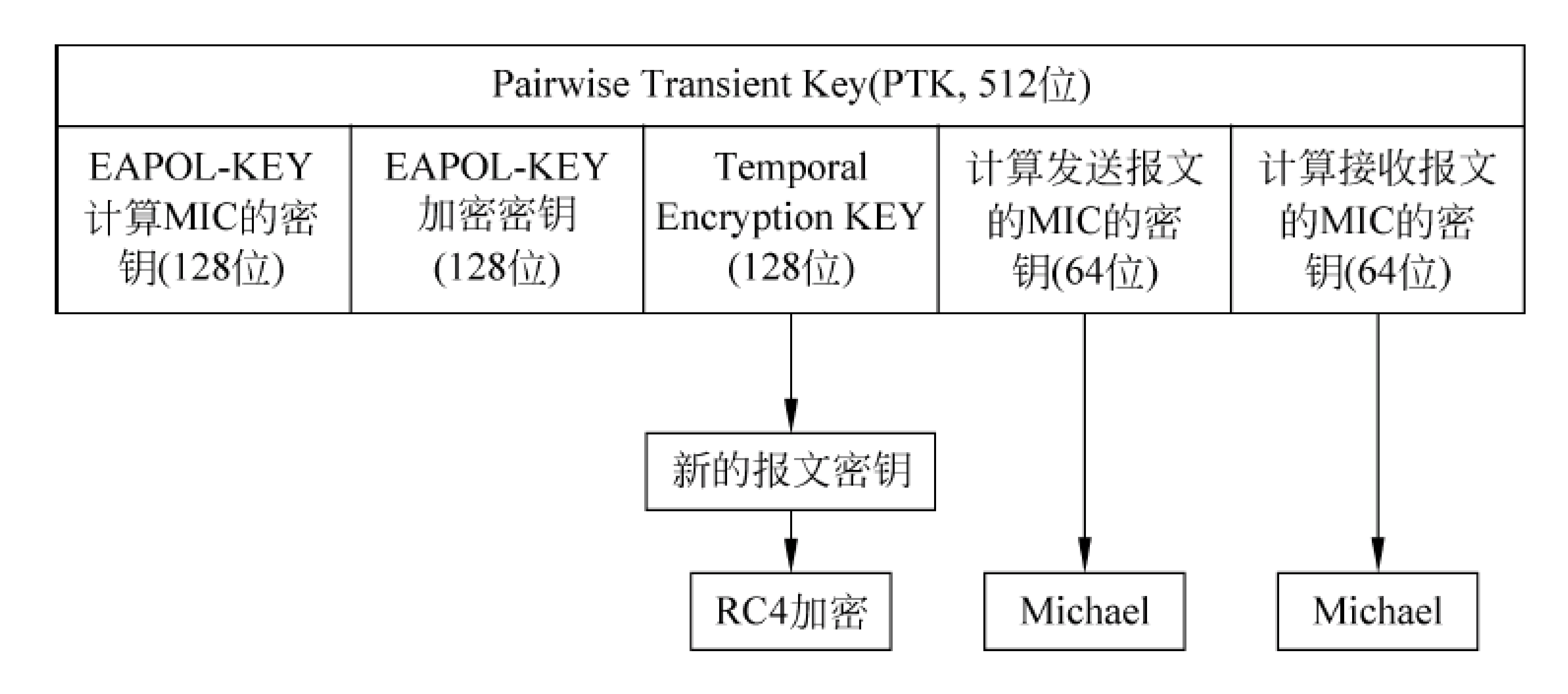


图 7-7 PTK 的产生过程

为了使现有的设备能够通过软件/固件升级实现 WPA，协议规定 AP 可以不采用 PTK 方式，而是利用下面将要描述的 GTK 作为 AP 向 STA 发送单播报文时的密钥。如果 AP 通知 STA 安装并使用 PTK，那么 STA 在向 AP 发送一个 EAPOL-KEY 相应报文后，再把相应的密钥安装到无线网卡中。

四次握手成功后，AP 要生成一个 256 位的 GTK(Group Transient Key)，GTK 是一组全局加密密钥，所有与该 AP 建立关联的 STA 均使用相同的 GTK，AP 用这个 GTK 来加密所有与它建立关联的 STA 的通信报文，STA 则使用这个 GTK 来解密由 AP 发送的报文并检验其 MIC。该密钥可以分解为三种不同用途的密钥，最前面的 128 位作为构造全局"每报文密钥"(per-packet encryption key)的基础密钥(base key)，后面的两个 64 位的密钥分别作为计算和检验 WPA 数据报文的 WIC 的密钥。AP 使用 EAPOI-KEY 加密密钥将 GTK 加密并发送给 STA，并指明该 GTK 是否允许 STA 用作发送报文所使用，STA 成功接收到该报文，将 GTK 解密后，向 AP 发送应答报文，并根据 AP 所指示的 Key Index 将其安装无线网卡的相应位置，如果 AP 使用 GTK 作为向某一 STA 单播传输的密钥，则该 STA 也需要使用 GTK 作为向 AP 发送单播报文的密钥。TK IP 并不直接使用由 PTK/GTK 分解出来的密钥作为加密报文的密钥，而是将该密钥作为基础密钥(3Base Key)，经过两个阶段的密钥混合过程，从而生成一个新的每一次报文传输都不一样的密钥，该密钥才是用做直接加密的密钥。通过这种方式可以进一步增强 WLAN 的安全性。密钥的生成方式如图 7-8 所示。

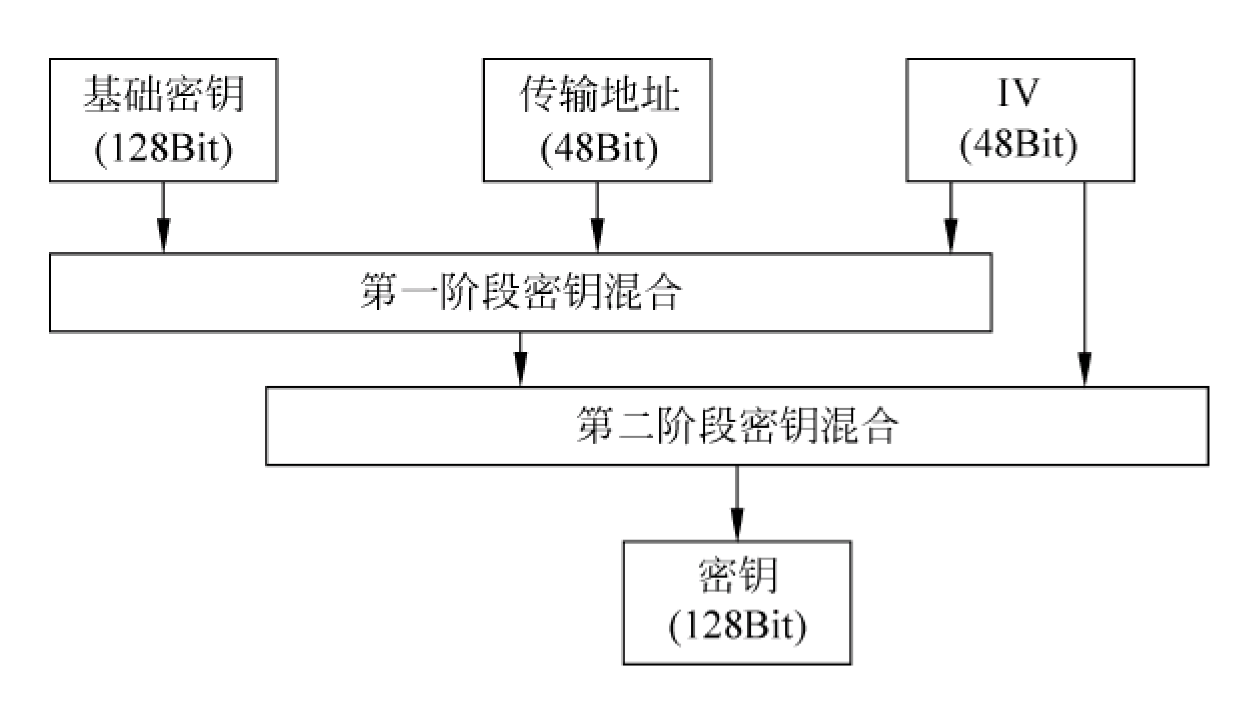


图 7-8 WPA 最终密钥生成过程

在 WPA 中，AP 支持 WPA 和 WEP 无线客户端的混合接入。但是在混合接入的时候，所有 WPA 客户端所使用的加密算法都得使用 WEP，这就降低了 WLAN 的整体安全性。尽管 WPA 在安全性方面相比 WEP 有了很大的改善和加强，但 Wi-Fi 联盟承认目前使用 KIP 的 WPA 只是一个临时的过渡性方案。

7.4.6 WAPI 协议

除了国际上的 IEEE 802.11i 和 WPA 安全标准之外，我国也发布了在 2003 年 12 月 1 日起强制执行的 WLAN 国家标准。此标准的一个重要组成部分就是由宽带无线 IP 标准工作组制定的新的安全机制——无线局域网鉴别和保密基础结构（WLAN Authentication and Privacy Infrastructure，WAPI）。WAPI 由 WAI（WLAN Authentication Infrastructure）和 WPI（WLAN Privacy Infrastructure）两部分组成，分别实现对用户身份的鉴别和对传输的数据加密。WAPI 采用公开密钥密码体制，利用证书来对 WLAN 系统中的 STA 和 AP 进行认证，其工作原理如图 7-9 所示。

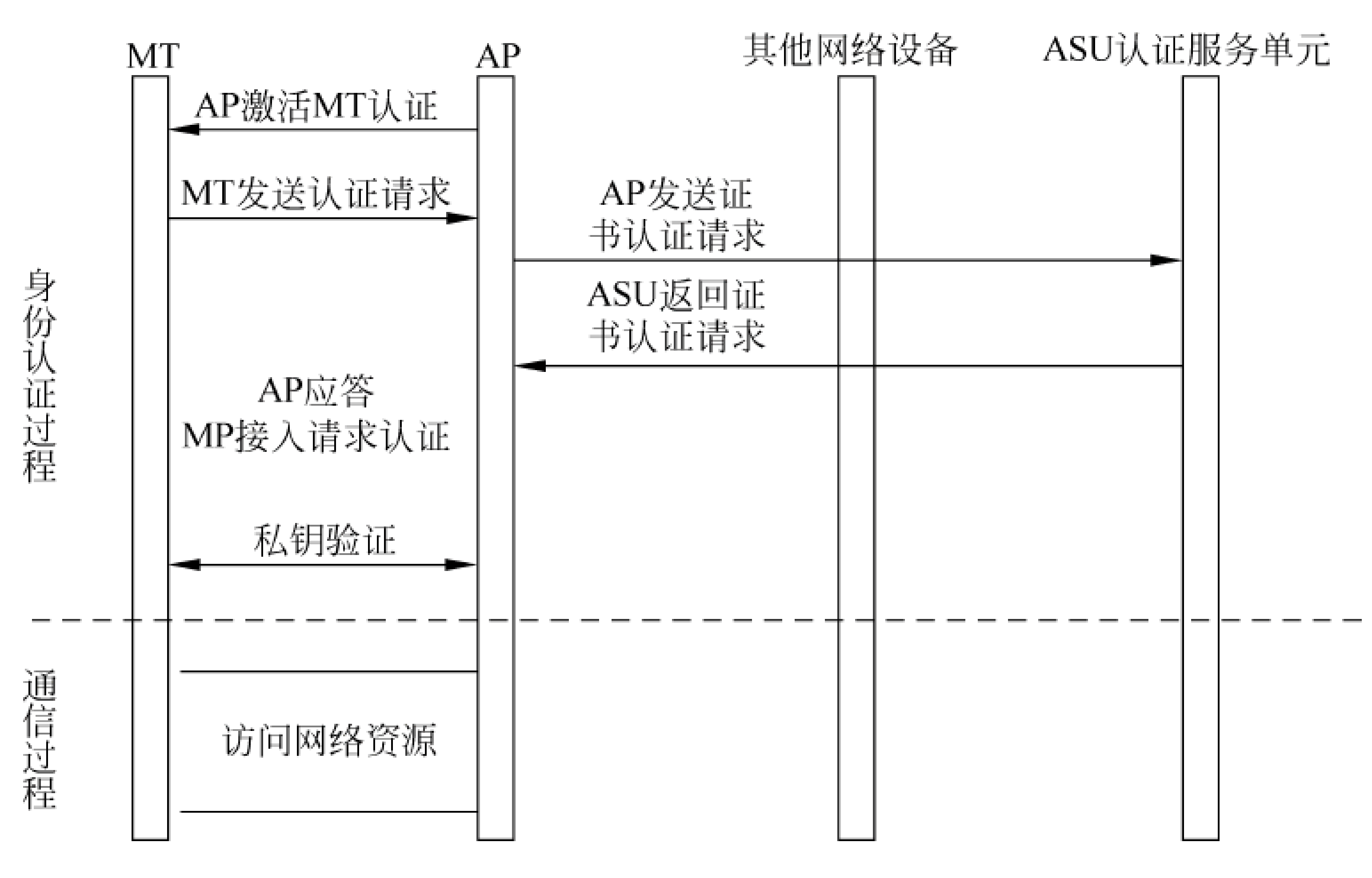


图 7-9 WAPI 工作原理

整个系统由移动终端 MT（Mobile Terminal），AP 和认证服务单元 ASU（Authentication Service Unit）组成，ASU 用于管理参与信息交换各方所需要的证书（包括证书的产生、颁发、吊销和更新）。证书里面包含有证书颁发者的公钥和签名以及证书持有者的公钥和签名（采用 WAPI 特有的椭圆曲线数字签名算法），是网络设备的数字身份凭证。WAPI 的工作原理如下：

（1）认证激活。当移动终端 MT 登录到 AP 时，由 AP 向 MT 发送认证激活以启动整个认证过程。

（2）接入认证请求。MT 向 AP 发送接入认证请求，即将 MT 证书与 MT 的当前系统时间发往 AP，其中系统时间称为接入认证请求时间。

（3）证书认证请求。AP 收到 MT 接入认证请求后，向 ASU 发送证书认证请求，即将 MT 证书、接入认证请求时间、AP 证书并利用 AP 的私钥对它们签名构成证书认证请求发送给 ASU。

(4) 证书认证响应。ASU 收到 AP 的证书认证请求后,验证 AP 的签名以及 AP 和 MT 证书的合法性。验证完毕后,ASU 将 MT 证书认证结果信息(包括 MT 证书、认证结果及 ASU 对它们的签名)、AP 证书认证结果信息(包括 AP 证书、认证结果、接入认证请求时间及 ASU 对它们的签名)构成证书认证响应报文发回给 AP。

(5) 接入认证响应。AP 对 ASU 返回的证书认证响应进行签名验证,得到 MT 证书的认证结果。AP 将 MT 证书、认证结果信息、AP 证书认证结果信息以及 AP 对它们的签名组成接入认证响应报文回送至 MT。MT 验证 ASU 的签名后,得到 AP 证书的认证结果。MT 根据应该认证结果决定是否接入该 AP。

(6) 私钥验证请求。AP 和 MT 都需要确认对方是否是证书的合法持有者,私钥验证请求包含实时产生的随机数,请求对方对其签名,以验证对方是否拥有该证书的私钥。该请求可由 AP 或 MT 发起。

(7) 私钥验证响应。包含对私钥验证请求中随机数据的签名,提供自己是证书合法持有者的证明。

(8) 至此 MT 和 AP 之间完成了证书认证过程。若认证成功,则 AP 允许 MT 接入,否则解除其登录。

由于采用了双向认证,所以不仅可以防止非法移动终端 MT 接入 AP 而访问网络并占用网络资源,而且可以防止移动终端 MT 登录至非法 AP 而造成信息泄漏。另外会话密钥并没有在信息上进行传输,因此就增强了其安全性。为了进一步提高通信的保密性,WAPI 还规定,在通信一段时间或者交换一定数量的数据之后,STA 和 AP 之间可以重新协商会话密钥。

WAPI 具有以下重要特点:

(1) 全新的高可靠性安全认证与保密体制,更可靠的链路层以下安全系统,完整的"用户—接入点"双向认证,集中式或分布集中式认证管理,证书—密钥双认证,灵活多样的证书管理与分发体制,可控的会话协商动态密钥,高强度的加密算法,可扩展或升级的全嵌入式认证与算法模块,支持带安全的越区切换。

(2) 支持 SNMP 网络管理,符合"国家商用密码管理条例"。

(3) WAPI 从应用模式上分为单点式和集中式两种,充分考虑了市场应用。单点式主要用于家庭和小型公司的小范围应用;集中式主要用于热点地区和大型企业。

7.5　小结

本章首先对无线局域网进行了概述,介绍了无线局域网的协议栈、组成、拓扑结构以及无线局域网的应用及发展趋势。接下来针对无线局域网存在的安全风险进行了分析,指出当前无线局域网面临的安全风险以及针对无线局域网的攻击手段;并展开了对安全需求分析,讨论了在审计、认证和保密这三方面的安全需求的情况。接下来的章节介绍了无线局域网的安全技术以及安全协议。无线局域网安全是一个不断改善和升级的过程,当前无线局域网所采用的安全技术和安全协议在实际使用中仍然存在一定的缺陷,对这方面的改进及研究方兴未艾。

参考文献

[1] Salli K T, Hamalainen T, et al. (1998). Security design for a new wireless local area network TUTWLAN. 3.

[2] Terry Lisa. Wireless LAN Worries. Supply Chain Systems, 2003, 23(2): 6-8.

[3] Bluck L, Vollbrecht J. PPP Extensible Authentication Protocol (EAP), http://www. ietf. org/rfc/rfc2284. txt.

[4] Rigney C. Remote Authentication Dial In User Service (RADIUS), http://www. ietf. org/rfc/rfc2865. txt.

[5] Borisov N, Goldberg I, Wagner D. Intercepting Mobile Communications: The insecurity of 802. 11. Proceedings of the Seventh Annual International Conference on Mobile Computing and Networking, 2001: 180-188.

[6] Fluhrer S, Mantin I, Shamir A. Weaknesses in the key scheduling algorithm of RC4. Eighth Annual Workshop on Selected Areas in Cryptography August 2001 available from http://www. drizzle. com/~aboba/IEEE/rc4_ksaproc. pdf.

[7] Aiesnort. http://airsnort. shmoo. com.

[8] Wepcrack http://sourceforge. net/projects/wepcrack.

[9] Wu T. A real world analysis of kerberos password security. In Proceedings of the 1999 Internet Society Network and Distributed System Security Symposium.

[10] Borisov N. Goldberg I and Wagner D. Intercepting mobile communications: The insecurity of 802. 11. Proceedings of the Seventh Annual International Conference on Mobile Computing and Networking, pages 180-188, 2001.

[11] The WiMAX Forum, http://www. wimaxforum. org/home.

CHAPTER 8

第8章　无线Mesh网络的安全

摘要：无线 Mesh 网络是一种多跳、具有自组织和自愈特点的网络，是近年来发展迅速的宽带无线通信网络，但它在发展的同时，又面临着许多安全问题。本章首先对无线 Mesh 网络进行了概述，然后对无线 Mesh 网络的安全风险和安全需求进行了分析，最后重点阐述了基于 MSA 协议的安全协议及相关技术。

关键字：无线 Mesh 网络、安全风险、安全需求、安全技术、安全协议。

8.1　无线 Mesh 网络概述

8.1.1　无线 Mesh 网络基本概念

目前主要观点认为，无线 Mesh 网络是一种多跳、具有自组织和自愈特点的宽带无线网络，无线 Mesh 网络是一种由无线路由器和终端设备组成的静态无线网络，是 Internet 的无线版本[1]。

按照体系结构划分，无线 Mesh 网络可以分为三种[2]：主干网结构、终端组网结构和混合结构。

1. 主干网结构

在无线 Mesh 网络主干网结构中，网络中的 MR(Mesh Router，网状网络客户端)互联构成了骨干网络，如图 8-1 所示。这些 MR 可以分为两种，一种是具有网关功能的，如图 8-1 所示中的 MR with Gateway/Bridge，它们负责连接终端节点，实现终端节点的网络接入，并且能够实现不同标准通信子网之间的互联，如无线局域网、传感器网络、蜂窝通信网等，同时部分具有网关功能的路由器还负责连接 Internet，使得网络中各个节点能够访问 Internet 资源。另一种路由器是不具备网关功能的，它们只负责数据的转发，如图 8-1 所示中的普通 MR 节点。

这种主干网结构的无线 Mesh 网络是目前应用较为广泛的一种体系结构，MR 一般被部署在屋顶或者较高建筑物上，其射频器件一般分为两类：其中发送半径较短的射频器件用于与终端用户连接；发送半径较大的射频器件(如方向天线)用于骨干节点之间的数据传输。

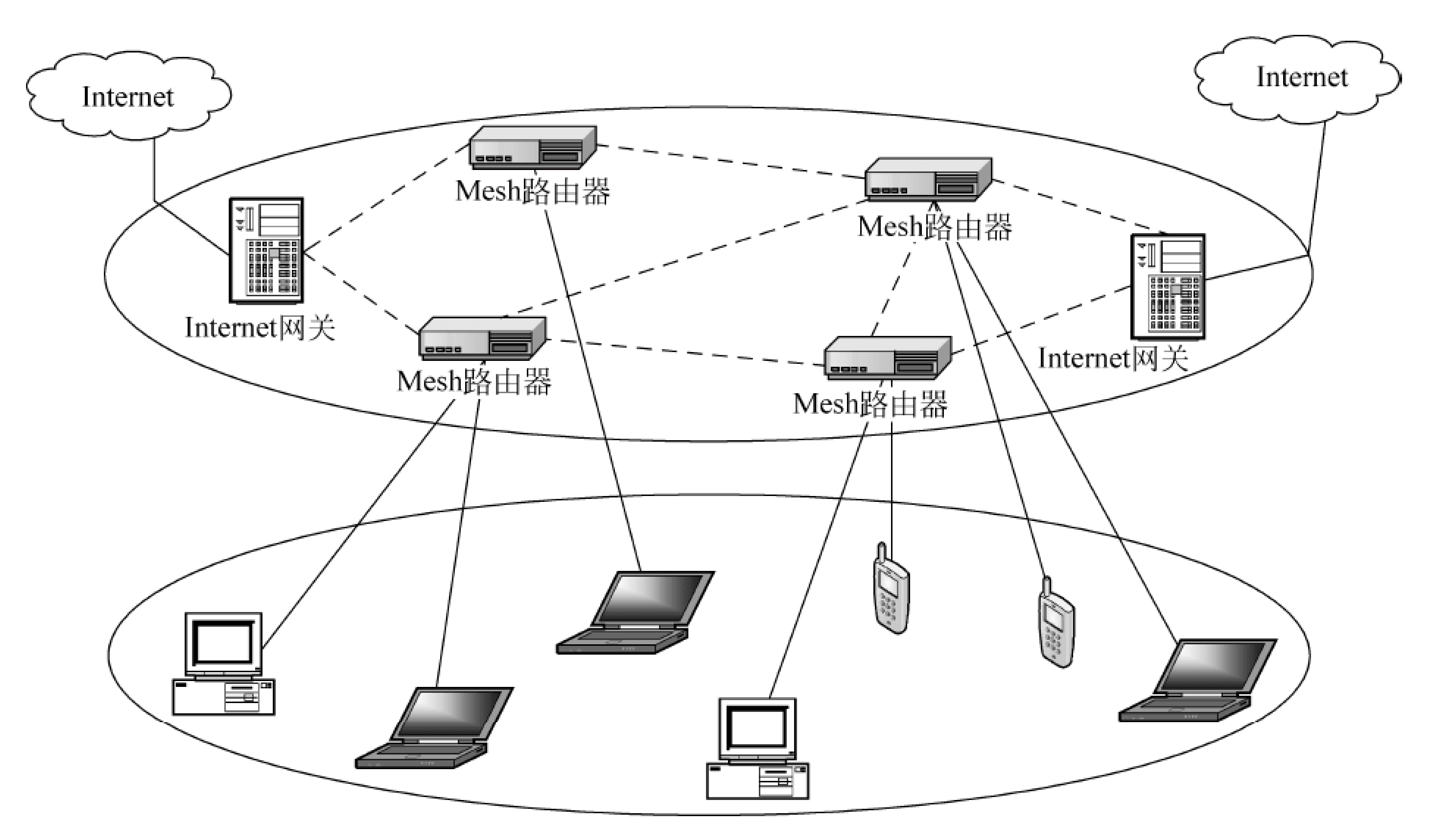


图 8-1　无线 Mesh 网络主干网结构

2. 终端自组网结构

终端自组网结构的无线 Mesh 网络由对等的终端节点组成,如图 8-2 所示。在该网络中,节点通过自组织,自配置方式组网,为终端用户提供端到端的服务。因此在这种结构中是不需要 MR 的。

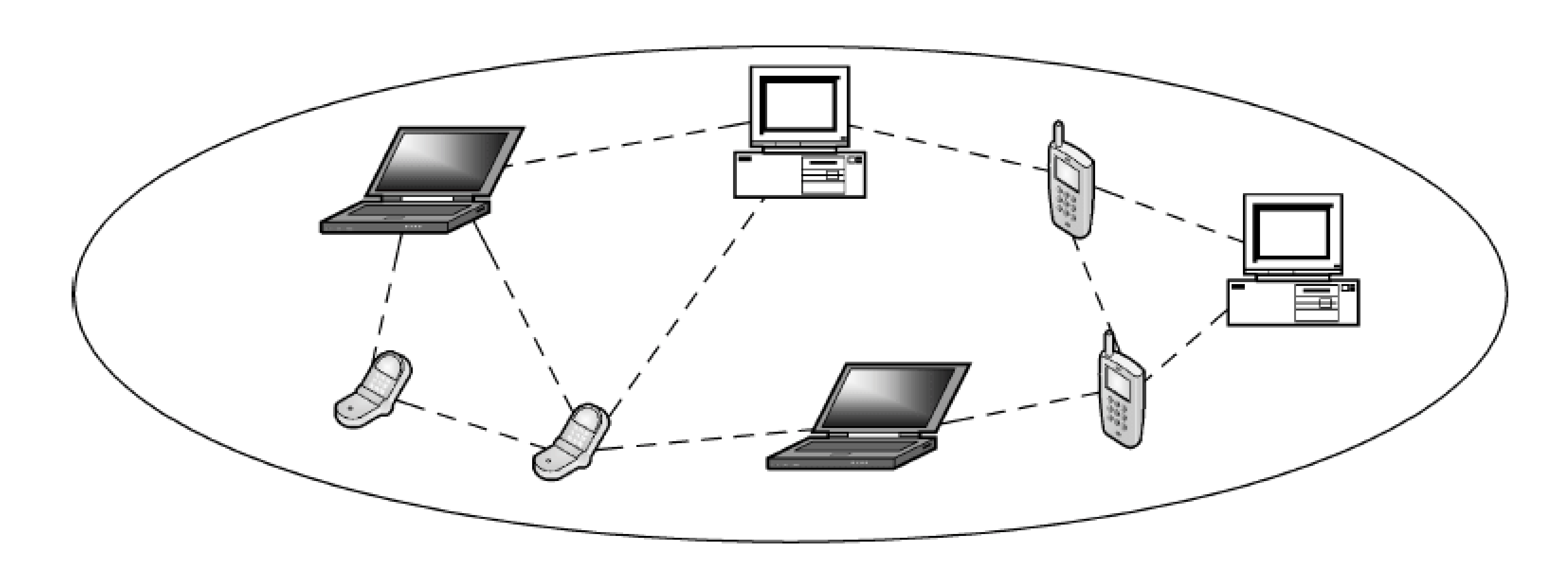


图 8-2　无线 Mesh 网络终端自组网结构

在终端自组网结构的无线 Mesh 网络中,当源节点发送数据包给目的节点时,数据包通过多跳的方式传送,中间节点负责路由和数据的转发,其功能相当于路由器。可以说,这种组网结构实际上等同于 Ad hoc 网络,但在移动性上仍有所不同。

3. 混合结构

如图 8-3 所示,混合结构的无线 Mesh 网络是主干网结构和终端自组网结构的有机结合。MC(Mesh Client,网状终端)可以通过 MR 实现网络的接入,也可以通过其他 MC 多跳转发实现接入。

混合结构的无线 Mesh 网络拥有更广的应用范围和更好的适应性。例如在紧急救援行动中,救援人员既可以用随身携带的 MC 临时组网,相互之间进行通信,又能够及时地将救援行动中的重要数据通过 Internet 发送到总部。

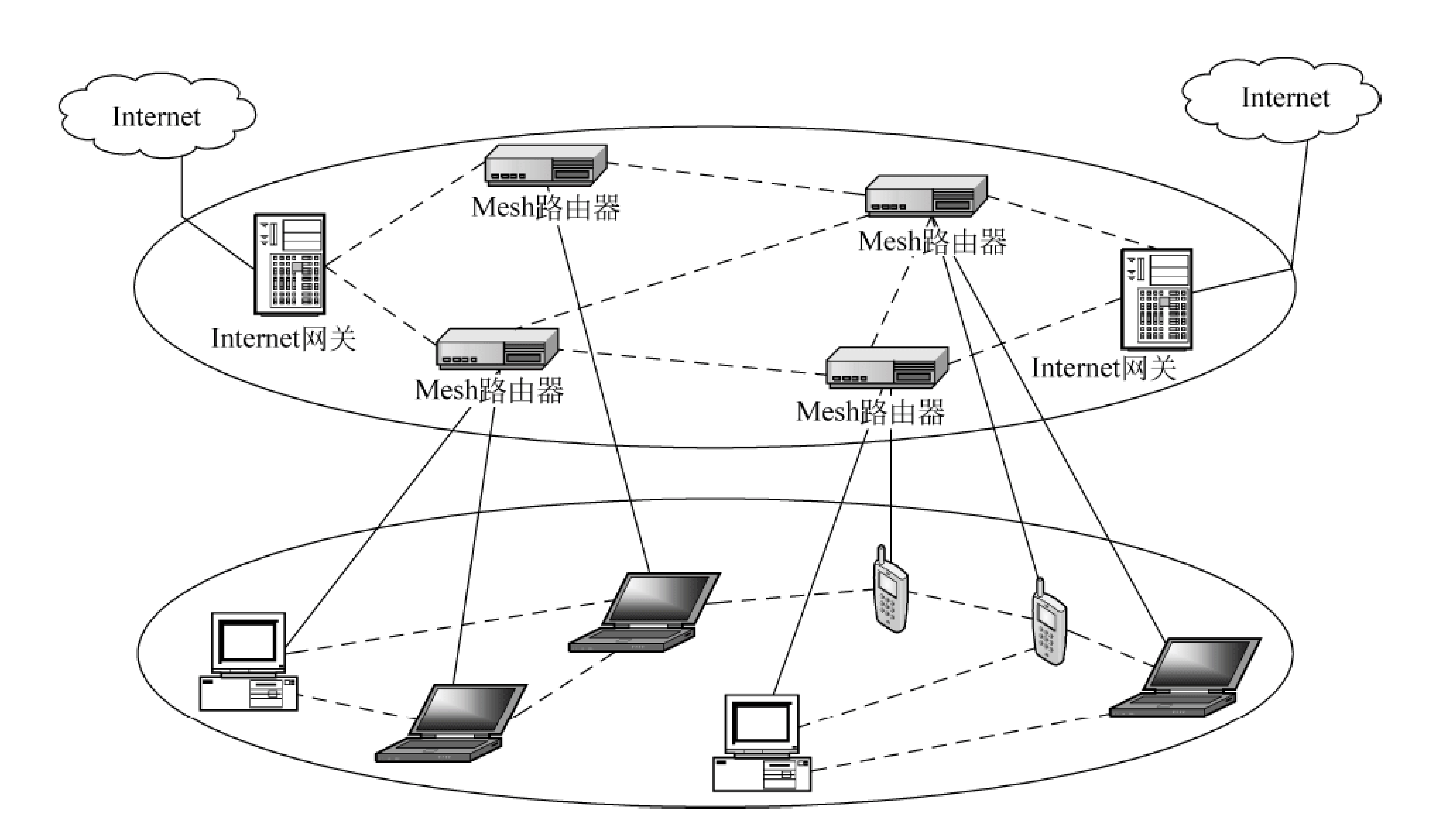


图 8-3　无线 Mesh 网络混合结构

8.1.2　无线 Mesh 网络标准化与产品化进展

1. 无线 Mesh 网络的标准化进展

鉴于无线 Mesh 网络的迅速发展，许多国际标准化组织都积极考虑在各种无线网络标准中加入对 Mesh 组网方式的支持。

IEEE 802.16 标准组在 2003 年 4 月颁布的 IEEE 802.16a 宽带无线城域网标准中设计了对点到多点和 Mesh 两种拓扑结构的支持。由于 WiMAX 技术可以在创建大范围无线回程网络中应用，故在其 Mesh 模式中应用了基于回程的 WiMAX[3]。

到目前为止，802.15.1～802.15.4 本质上均不能直接支持网状网络结构，而只是 P2MP(Point to Multiple Point，点到多点)方式下的微微网结构，但散射网已经有了无线 Mesh 网络的雏形[7]。正在制定过程的 802.15.5 标准继承了 802.15.1～802.15.4 的基本思想，且完全支持网状结构和移动性管理，其目标是利用短距离、低成本的设备以 Mesh 组网的方式去覆盖一个较大的环境，如客厅、校园、医院等。

传统的 802.11 MAC 层协议的固有属性并不支持网状连接，使得网络性能在多跳情况下很差。为此，IEEE 成立 802.11s 子工作组，制定标准化的扩展服务集(ESS)[4,5]，专门为无线 Mesh 网络定义 MAC/物理层协议及其上层的无线分布式系统的协议，以便使无线局域网的多个 AP 之间能够通过自动配置拓扑来组网。

此外，国际电信联盟(ITU)、3GPP 以及 IETF 等机构也将无线 Mesh 网络纳入到工作计划中，这些组织结构的标准化工作必将进一步推动无线 Mesh 网络技术的研究、应用和推广。

2. 无线 Mesh 网络的商业产品化进展

作为一种新兴的网络形态，无线 Mesh 网络可应用于很多领域，其应用方式主要有无线

接入、无线传输和简单组网。面对无线 Mesh 网络的迅速发展，一些致力于其开发和应用的公司，如美国的 Mesh 网络、加拿大的 Nortel 等，都推出了自己的无线 Mesh 网络设备和相应的组网技术，并已应用这些设备和技术成功地解决了一些热点地区的无线接入问题，其中比较成功的解决方案主要有：俄勒冈警察局设计的应急通信方案（Mesh 网络）、英格兰 Portal 智能交通系统提供的应用方案、Nortel 设计的移动城市应用方案等。

除了 Mesh 网络和 Nortel，目前的无线 Mesh 设备和解决方案的提供者主要还有 Skypilot、Tropos、BelAir、Firetide 和 RadiantNetwork 等公司。对于这 7 家公司的产品比较如表 8-1 所示，从物理层/MAC 层技术来说，大多数产品选择了直接沿用 802.11 系列的物理层和 MAC 层规范，工作频段基本上为 2.4/5.8GHz；从路由协议来看，二层路由和三层路由比例相当；产品的覆盖范围取决于该产品的用途，一般用于主干网和室外的节点覆盖范围远大于室内的节点覆盖范围；另外，可以发现目前商业产品对 QoS 的支持并不是很普遍，只有一半左右的产品支持。

表 8-1 主要的无线 Mesh 网络商业产品比较[6]

产品名称	公司	用户速率(Mb/s)	物理层技术	频段(GHz)	节点范围	MAC 协议	路由位置	QoS 支持
Skypilot System	Skypilot（美国）	>3	IEEE 802.11a	5.8	32/6.4km (LOS/NLOS)	Skypilot 同步协议	第 2 层	Diff Serv
Tropos 5110/3110	Tropos（美国）	0.512-1	IEEE 802.11b	2.4	4.15km/290m（室外/室内）	802.11 MAC	第 2/3 层 PWRQ	/
MEA	Mesh Networks（美国）	1.5-6	QDMA	2.4	1.6km (NLOS)	多信道 MAC	第 2 层 MSR	Diff Serv
BelAir 200/100	BelAir（加拿大）	>1	IEEE 802.11b/g/a	2.4/5.8	500/100m（主干/用户）	802.11 MAC	第 2/3 层	QoS
Wireless AP220	Nortel（加拿大）	>1	IEEE 802.11b/g/a	2.4/5.8	1km/100m（主干/用户）	802.11 MAC	第 2 层	/
Hotpoint 1000	Firetide（美国）	>1	IEEE 802.11b	2.4	3km/200m（室外/室内）	802.11 MAC	第 3 层	/
MESH WORKS	Radiant-Network（英国）	>4	QPSK/QAM	5.8	2km（室外）	/	ATM	ATM QoS

8.1.3 无线网状网络与现有无线技术比较

现有的网络技术层出不穷，无线网状网络能够在其中成为新一轮的研究热点是由它本身的特点和未来网络发展的方向所决定的。表 8-2 对无线网状网络、移动 Ad hoc 网络以及蜂窝网络做了比较，无线网状网络的优缺点一目了然。

由于无线网状网络是基于 Ad hoc 网络发展而来的，所以两者的密切关系也值得探究，两者在内部系统上保持了相当的一致性，二者具有以下共同点：

（1）网络的自组性。人们对于互联需求的不断提升，任何时间、任何地点、任何方式的连接需求要求网络必须能自动地面对所有可能的问题。

表 8-2 无线网状网络与其他网络比较

比 较 项	无线网状网络	移动 Ad hoc	蜂 窝 网 络
拓扑结构	多点到多点(网状)	动态拓扑	点到多点
控制方式	分布式控制	分布式控制	集中式控制
覆盖范围	城域覆盖	局部范围	覆盖广大地区
设计目的	用户接入为主	用户间通信为主	接入和通信同时
容纳用户数	多	较少	非常多

(2) 有限的无线传输带宽,无线技术相对有线技术易受干扰的特性,使得在传输速率上可能会形成较大的差异,而且多个会话同时占用信道导致速率波动更加明显。

(3) 多跳通信方式,这一点作为和传统集中式网络结构最大的不同点,要求各接入 AP 的地位互相平等,可以在任何情况下选择最优路径,而不会出现竞争的延迟,同时降低节点功率达到节能效果。

无线网状网络与移动 Ad hoc 网络最大的区别有以下两点:

(1) 无线网状网络比移动 Ad hoc 网络的要低,网络拓扑变化没有那么频繁,在一段时间内可以认为是静止的。

(2) 移动 Ad hoc 网络中的业务流量主要是来源于各个终端之间的通信,而无线网状网络中的业务主要是和 Internet 通信。

8.2 安全风险与安全需求

IEEE 802.11(WLAN)、802.15(WPAN)和 802.16(WMAN)是当今主要标准,故 WMNs 的部署也基本可分为三种:802.11 Mesh、802.15 Mesh 和 802.16 Mesh。由于无线局域网 Mesh 网络应用最为广泛,所以本章重点讨论无线局域网 Mesh 网络的安全问题。

8.2.1 无线局域网 Mesh 网络常见的安全威胁

在人们享受无线网络所带来的便捷性的同时,伴随而来的是无线网络所具有的安全威胁[7],WLAN Mesh 网络同样如此,而且 Mesh 这种特性还会招致一些额外的攻击和威胁,我们将在下面分别介绍。

1. 窃听、监听和截取

随着网络技术的不断发展,窃听(sniffing)已经从网络通信分析工具演变成了网络攻击者手中的利器。广义的窃听涵盖了从遍历网络映射目录到密码获取及捕获未加密数据,各种形式的获取未授权信息行为。

这种偷听流经网络的计算机通信的电子形式,就是窃听。在有线网络中,最先部署设置阶段允许网络中的每一台机器看到其他机器通信,以使转发器和集线器可以将整个网络连接在一起,而对于攻击者来说,只要连接到可以看到整个网络通信的节点上,就可以轻松获得整个网络的所有通信。而无线网络的功能特性恰恰与最先部署阶段的集线器和转发器非常相似。对于监听网络的用户来说,流经整个通信网络的信息都是可见的,监听者甚至不需要连接到网络中仍然可以监听到那些未加密的数据。对于无线 Mesh 网络来说,这种窃听

攻击更加容易和危险,因为在任何一个 Mesh 节点都在网络的路由结构中起着作用,而且这种作用和重要性随时可变也不能预先确定,因此对于安管人员来说无法预先确定对于网络各个部分的安全保护机制要求多高。

保护无线网络用户免受攻击者窃听的最有效途径就是对所有在可能被监听的区域传输的数据包进行加密,在 Mesh 网络中,也要一套完善而易于使用的加密机制。

2. 欺骗和非授权访问

最常见的欺骗攻击手段是 IP 欺骗。使用 IP 欺骗的攻击者甚至不需要详细了解 TCP/IP 的指示,现有的模块化窗体界面操作工具让初学者同样可以进行极具威胁的欺骗攻击。

欺骗攻击对 WMN 的各个层次构成严重的威胁。如果没有足够安全的用户身份验证,在网络层,泄密的节点可以冒充其他受信任的节点攻击网络,在网络管理的范围内,攻击者可以获得超级用户权限,从而访问配置系统,在服务层次,一个恶意的用户甚至不需要适当的证书就可以拥有经过授权的公钥,这对某些安全需求来说几乎是毁灭性的。

成功假冒造成的损失非常严重。一个恶意用户可以冒充任何友好节点,导致整个网络拓扑结构混乱到不可识别,并对其他节点或者服务造成永久性毁坏。对付欺骗和非授权攻击,比较有效的方式是外部身份验证资源,这样可以防止非授权用户访问无线网络及其连接的资源[8]。

3. 网络接管与篡改

攻击者可以通过多种技术接管无线网络或者接管会话过程。因为在很多情况下,对网络本身甚至管理员也会很难区分出攻击者和合法用户的不同行为。

在 WLAN Mesh 网络中,一种极具危险性的攻击手段就是使用假冒 Mesh 接入节点 MAP[9],攻击者可以部署一个发生强度足够大的 MAP,这可能导致终端节点无法辨认 MAP 的真伪,并使用恶意 MAP,利用这一点,攻击者可以接收到身份认证请求和来自终端节点与密钥相关的信息。

还有一种攻击则是通过 ARP,可以伪造与 MAP 的连接,假扮目标主机,所有试图在主机上部署 SSH 的用户将被连接到假冒的主机上,在用户进行身份验证时,攻击者可以接受到用户密码信息,然后一个经验丰富的攻击者可以通过转发密钥给真正的目标节点,以掩盖其攻击行为不被用户发现。

4. 拒绝服务攻击

拒绝服务攻击会导致节点无法对其他合法的节点或者中断提供所需的正常服务。在物理层和 MAC 层,攻击者通过拥塞无线信道干扰通信;在网络层,攻击者破坏路由信息,使网络无法互连;在更高层,攻击者可以通过伪造使高层服务陷入无验证性可言的境地。

拒绝服务攻击的严重性取决于 WMN 的实际应用环境。在 WMN 中,使中心资源溢出的拒绝服务攻击威胁有限;相反,分布式的拒绝服务攻击威胁更大,如果攻击者计算能力足够强,带宽足够大,WMN 很容易阻塞甚至崩溃。

然而,对 WMN 更严重的威胁是被占领的 MP 可能会重新配置全部或者部分路由信息,这在前面已经提及过,从而造成网络阻塞,具体来说就是一个被占领的 MP 欺骗相连 MP,从而导致错误路由信息放射状外延,从而阻止其他节点获得已改变的网络拓扑信息,在最糟糕的情况下,攻击者通过改变路由协议,使整个网络置于其控制下。

5. 物理攻击

最简单的攻击手段往往也就最直接有效，这个道理同样适用于黑客，许多安全管理员可能认为笔记本电脑，PDA，Web 电话等硬件设备丢失并不会对网络本身造成影响，但黑客认为，任何具有 Web 功能的设备都是非常有价值的，因为它可以让他们获得重要的用户信息，身份认证信息和所要入侵的网络的必须信息。

当失窃的网络设备中含有被盗用户的网络访问方式的信息时，罪犯就可以通过无线网络访问受限信息，如果被盗的 Laptop 中包含 PGP 密钥环信息，很显然利用伪造签名可以进行一些烈的社交工程攻击，以及破解系统中的加密文件和数据流。这实际上和破解密钥环从而能达到的攻击性是几乎一致的，而所需要的工作量却有天壤之别。

WMN 中，MP 的数量要远远多于集中模式下的 AP，每个 Mesh 节点肩负着重要作用而又位置分散，网络安全管理员很难保护到所有的 MP，而使攻击者有可乘之机。

8.2.2　WLAN Mesh 网络安全需求

对于 8.2.1 节中列举的 WLAN Mesh 网络攻击方式，我们需要一些专门的安全技术来防范，这就是本节将要介绍的 WLAN Mesh 网络安全需求[10]，这些安全技术在其他无线网络中也常常被应用，为了适用于 Mesh 网络，有些地方还是要做相应的修改。下面分别介绍。

1. 身份认证

身份认证可以有效防止假冒用户和假冒网络的安全威胁。其他的安全机制，如接入控制和数据加密等往往构建在身份真实性的基础之上，因此身份认证可称得上网络安全的第一道屏障。身份假冒的威胁是两方面的，一方面攻击者可以假冒合法用户骗取网络接入授权，另一方面攻击者可以部署假 MAP 冒充网络。因此单单网络对用户身份的认证是不够的，必须实现网络和用户的双向认证。

2. 授权和接入控制

认证(authentication)和授权(authorization)是两个有联系但又相互区别的概念，认证解决“你是谁”的问题，授权决定“你能做什么”。虽然大部分情况下，通过认证也就意味着获得授权；但严格说起来，两者并不是一回事。接入控制是为了根据用户的身份、角色以及预订服务等对用户能访问的资源进行限制，常见的接入控制机制如在路由器上创建 ACL 列表等。在 WMN 中，路由信息将在分布式的 MP 上分散分布，所以如何处理这些授权和接入控制数据库将是另一种方式。

3. 通信数据的保密性

保密性保证通信数据只被希望的接收方看到，为防止恶意者的窃听，应该采用强壮的加密算法对通信数据进行加密封装。在传统的 WLAN 集中模式中，接入移动终端可以和 AP 之间建立加密密码，但是在 WMN 中，分布式的 MP 之间如果都采用同一的密码，将极大地降低安全性，一旦有一处被攻破，将导致整个 Mesh 网络的数据泄露；如果两两 MP 之间采用不用的加密密码，那么密码的数量将随着 Mesh 网络的规模而急剧增加，不仅不便于存储，也不便于管理。所以，如何制定一套 WMN 环境中的密码体系，如何管理密码也需要特

殊的方式。

4. 数据和信令的完整性

对报文数据进行完整性保护是保证数据真实性的必要手段,单纯的加密并不能防止数据被篡改。不仅用户通信数据,网络控制信令也要防止攻击者的恶意篡改。篡改信令可能给网络造成严重的后果。同保密性一样,对 WMN 分布式的环境也需要数据传输过程中完整性密码进行特别管理。

5. 数据的顺序性

报文的顺序性属于报文的真实性的一个方面,它是指接收方必须可以确定收到的报文确实按照发送方原有的顺序,没有被删除、插入、重放和重新排列等。

在 WMN 中,MP 之间可以相互通信,这种复杂的通信关系必将导致通信的数据的复杂性,对顺序性的要求也会激增,如何管理好 MP 的通信模型,将是一项艰巨的任务。

6. 密钥管理

密钥管理是由于数据的保密性和完整性衍生出来的安全需求。广义的密钥管理概念包括密钥的产生、分配、传递、保存、恢复、销毁等;狭义的密钥管理指的是密钥协商。对于具有多个节点的网络系统来说,有效的密钥管理机制是必不可少的。首先,对于稍大一点的网络,要对所有的用户采用手工的方法配置密钥显得不切实际;其次,如果没有集中自动的密钥更新机制,人们会因为害怕麻烦而不再更新密钥,这样由于密钥不能及时老化更新而威胁到通信数据的安全性。

8.3 无线局域网 Mesh 网络特有的安全问题

前面列举了 WLAN Mesh 网络最常见的安全威胁和由此产生的安全需求,这些在其他的无线网络中也比较常见,下面将分析一些 WLAN Mesh 网络特有的安全问题[11],对于 WLAN 的 Mesh 模式,它所遭受的攻击如图 8-4 所示。

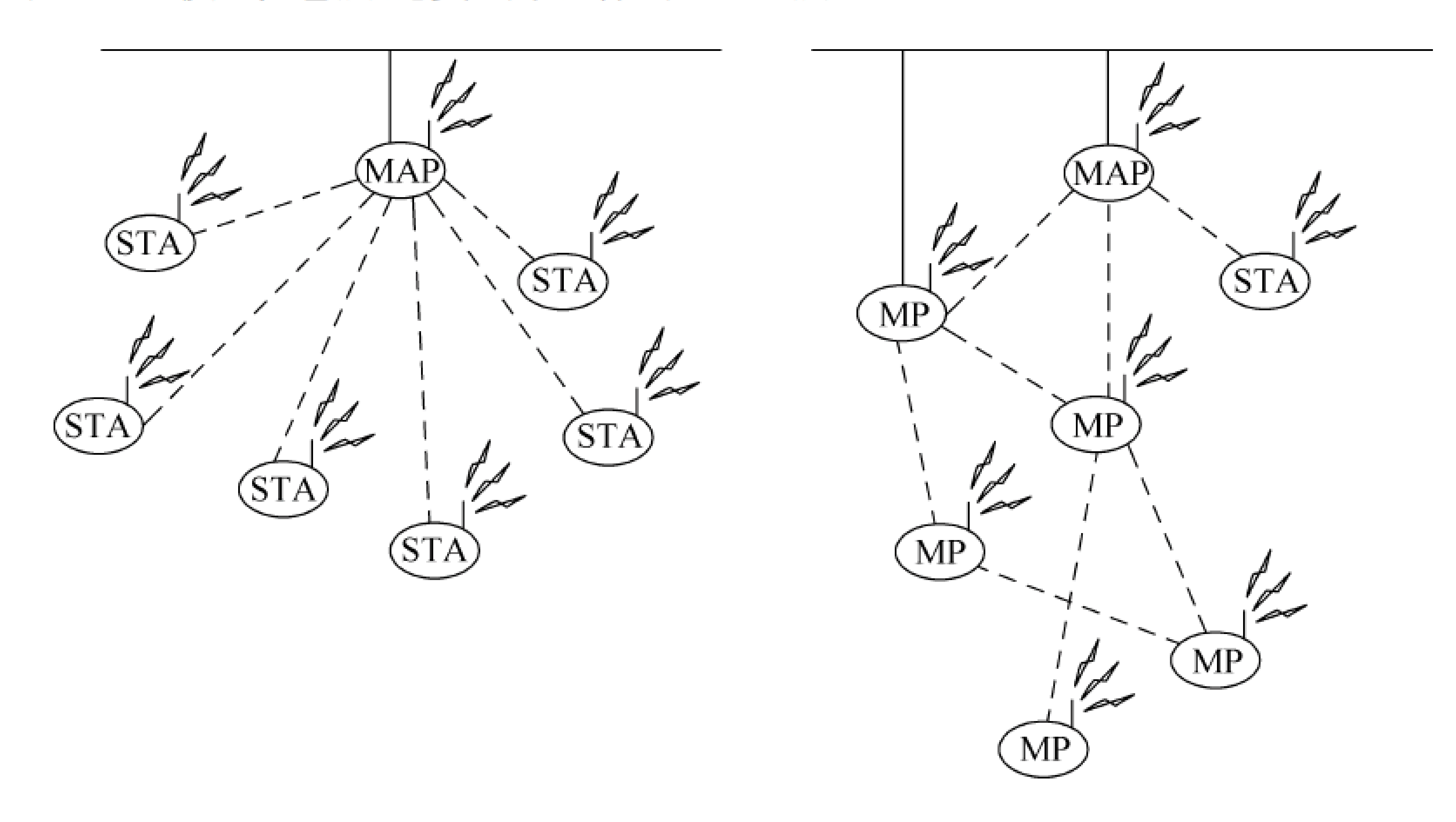


图 8-4 Mesh 网络攻击示意图

其安全威胁分为STA被攻击、MP被攻击和MAP被攻击。其中STA被攻击与WLAN基础设施结构中威胁相同，Mesh模式特有的威胁在于MP和MAP被攻击，针对这两种Mesh节点的安全问题又以多种形式存在，下面将分别介绍。

8.3.1　决策分散

WMN中的决策往往是分散的，许多WMN算法依赖于所有节点的参与合作。集中权力结构的缺乏意味着攻击者能够利用这些弱点实施新型的，用于破坏写作的算法进行攻击。与传统的集中式WLAN相比，Mesh网络的每一个MP都处以一个相对来说比较重要的地位[12]，对于整体网络的认证和路由功能发挥着作用，这种相对的平等性将导致决策的分散性。

对于这种安全问题，我们可以考虑设立一些具有额外决策功能的MP节点，这些MP并不需要如传统WLAN中的AP的复杂功能，又比普通的MP多了一些附加功能。

8.3.2　Mesh网络认证的问题

传统的WLAN，用户节点的接入可以通过AP的认证，但是对于Mesh网络，没有一个相对集中的认证节点存在，而且，由于Mesh网络中节点的相互平等性，要实现双向认证将意味着实现一个全网络的连接边数，才能保证任何一个Mesh节点对于其他所有可能连接的Mesh节点都是被认证的，但是随着Mesh节点的数量增长，这种认证数量将快速增长，如果MP的数量为n，则需要认证2的$n-1$次方次。当n大于一定数目的时候，认证次数的激增将导致网络通信流量的暴涨，这是不现实的，所以必须要有独特的认证解决方案来实现MP节点之间的双向认证。

另外，在新设备加入Mesh网络时，需要某种机制保证新设备可正确识别其他Mesh成员的角色。传统情况下，设备都是和MAS(Mesh Authentication Server)通信，或者当AS(Authentication Server)处于外部网络时，设备和端口点PP(Portal Point)通信。在802.11 Mesh中，设备和什么节点认证，或者还有可能和MAS或PP以外的其他设备通信，还是一个问题。

对于这种问题，我们考虑采用诸如EAP等方式的认证，同时设立相对集中的认证点和认证服务器，这还要和传统的WLAN所不同，可以解决双向认证和保持高效认证的统一方式。

8.3.3　多跳路由安全

由于无线网络的覆盖范围有限，一个无法直接接入到路由器的MP可以借助其他的MP转发进行数据通信，可在没有固定设施的情况下，通过移动节点间相互协作保持网络互联，拓展了移动通信的网络范围。这直接导致网络存在窃听攻击和终端服务的隐患。

WMN的路由也呈现特别的脆弱性[13]，在有线网络中，可在路由器和网关处进行特别的保护，而WMN则不同，劫持WMN节点的供给者能够通过散播错误的路由信息使整个网络瘫痪。更严重的是由于失密节点给出的错误路由信息对来自所有节点的信息都产生影响。

在 Mesh 网络中，由于路由的脆弱性，拒绝服务攻击变得更加容易，形式也有所改变，Mesh 网络中任何一个 MP 节点都可以成为“攻击者”，通过路由手段，实现一些欺骗，比如说“黑洞”：攻击者发送一些假冒的包，模拟一个有效的 Mesh 节点，再丢弃包；还有“灰洞”和“虫洞”等。

因此，这种由于 MP 的脆弱性、移动性导致的路由威胁在 Mesh 网络被放大了，需要一个安全的路由协议来保证 Mesh 网络的通信畅通。

8.3.4 自组织与资源分配问题

WMN 的节点一般是静止的，但在出现拓扑变化即新节点加入或旧节点退出、链路干扰的情况下可能会发生变化。此时，WMN 自组织特性[14]就会自动维护网络的可用性。在此过程中，一个经验丰富的黑客很容易抓住这个机会。采用假冒攻击，拒绝服务攻击等手段实现网络入侵或导致网络瘫痪。对于自组织的 WMN 来说，最严重的威胁是破坏网络的可用性，如果一个假冒 MP 或者泄密的 MP 可能会广播虚假路由信息，会导致 WMN 面临两种威胁，一种威胁是所有的通信量被定向到攻击者所在的 MP，攻击者及此刻已窃听，篡改，直接导致整个网络通信中断。毫无疑问，此时网络传输速度也会由于通信量过于集中而导致网络传输速率大幅下降。攻击者的后续行为极有可能控制整个网络。另一种威胁是被入侵 MP 利用发送虚假路由信息欺骗与之有联系的 MP，会导致错误路由信息的破坏作用的放射状增幅，直接导致整个网络崩溃。

由于无线路由器距离 Internet 接入点有近有远，远离 Internet 接入点的节点有可能获得很小的带宽，所以设计合理的协议来保证节点间公平[15]是很重要的，对公平性的保护也带来了新的挑战。

这两个问题也可以划归到安全路由的范围，通过建立一个安全的、健壮的、自适应的路由，实现所有的 MP 的资源公平性与自组织性[16]。

8.3.5 角色定义与切换

两个 MP 之间如何建立安全连接，它们又如何识别对方，如何确保它们能够正确识别对方为 MP 而不是 STA 或 NFMP(Non Forwarding Mesh Point)，这就需要一种安全机制保证 MP 之间可以安全正确的识别对方。

尽管授权节点可以承担多种角色，但是完全没有必要让节点始终扮演多种角色，将带来不必要的能量损耗。例如，一个 MAP 节点可以在进入到节能模式之前，切换到另一个角色。这就需要某种安全机制已保证各种授权角色之间的安全切换。

在 802.1x 标准中指出若 AS 对认证方来说是不可达的，那么 802.1x 认证将失败。故当 AS 在 Mesh 外部网络时，Mesh 接入点 MAP(Mesh Access Point)的角色将不能得到保证。在这种情况下，如果让 PP 承担 AS 的角色，又会引入新的问题：即是否要在所有的 PP 上加入 AS 代理。此外，还会有些非 PP 节点既属于 Mesh 网络，同时又属于外部网络，对他们的角色和功能分配将变得复杂。

必须要建立一个定义这种角色分配和转换的协议，实现动态的安全的角色切换。

8.4　基于 MSA 协议的安全协议及相关技术

8.4.1　基本概念

为了保护多跳、自组织的 WLAN Mesh 网络的链路安全，IEEE 802.11 TGs 工作组专门提出了一个称为 Mesh 安全关联（Mesh Security Association，MSA）的安全方案[17]。MSA 是一个安全框架，与 WLAN 中所应用的 IEEE 802.11i 方案相比，它使用了新的密钥体系，并规定了一系列新的认证协议建立并运用这一密钥体系。其最终目的是对每个 MP 节点在接入网络前进行可靠的身份认证，并为每一条链路两端的 MP 节点协商一套共同的密码算法及密钥，保护该链路上传输数据的机密性和完整性。

1. 密钥持有者 Key Holders

MSA 架构将参与安全交互的 MP 节点分成 3 种角色：Mesh 密钥分发者（Mesh Key Distributor，MKD）、Mesh 认证者（Mesh Authenticator，MA）和 Supplicant MP。其中 MKD 和 MA 合称为密钥持有者 Key Holders，用以取代原先单纯的 AS-Authenticator 模式。Supplicant MP 指在 802.1x/EAP 认证中作为 Supplicant 方的节点，一般为希望通过身份认证加入 Mesh 网络的 Candidate MP 节点。

MKD 作为 AS 在 Mesh 网络中的代理人，主要负责主密钥的生成和分发以及确认 MA 的资格。MKD 的引入就不需要每个 MA 节点都维护到 AS 的安全链路，保证 Mesh 网络内部的自组织特性。默认 MKD 与 AS 间维护着一条安全路径，保证它们之间传输的密钥信息的安全。MKD 与 MA 之间可以经过无线多跳路径，因此需要建立一支密钥体系以保证其间传输数据的安全。一个 MKD 定义了一个 MKD 域，用一个 MKDD-ID 标识。一个 MKD 域也是一个狭义的 Mesh 网络[18,19]，一个 MP 节点在某一时刻只能属于一个 MKD 域。

MA 是具备为 Candidate MP 节点提供认证服务资格的节点，它能够建立并维护一条通往 MKD 的安全链路以保证经其转发的 Candidate MP 的认证信息的安全。

MKD 的角色是人工指定的，一般位于 Mesh 网络与外部网络接口的网关节点上，需要与外部 AS 服务器相连。而 MA 和 Supplicant MP 的角色则是自适应的，一个 MP 节点随应用场景的不同所担当的角色也会有变化。只有在初始 MSA 认证中的角色选择阶段才能确定各自的相对角色。

2. MSA 协议集

MSA 架构定义了一系列协议完成用户认证与密钥体系的建立及维护工作。以下 MSA 协议是该架构的主体。

（1）初始 MSA 认证协议（Initial MSA Authentication）[20]：用于安全地建立 MP 对间的链路，并且在需要时实现安全认证和用于保护后续链路的密钥体系结构的建立。

（2）简化 MSA 握手协议（Abbreviated Handshake）[21]：用于使用已建立的密钥体系结构中存储的共享密钥来安全地建立 MP 对间的链路。

MSA Key Holder Communication 由 4 个协议组成[22]。

（1）Key Holder 安全握手协议（Mesh Key Holder Security Handshake）：在 MA 和

MKD间建立通信和安全关联。

(2) Key Holder安全解除协议(Mesh Key Holder Security Teardown)[18]：用于Mesh Key Holder间协商删除已建立的安全关联。

(3) Mesh密钥传输协议(Mesh Key Transport Protocol)：实现MA和MKD间密钥分配和管理。

(4) Mesh EAP消息传输协议(Mesh EAP Message Transport Protocol)：描述MA和MKD间EAP报文传输机制。

3. 安全关联

安全关联(SA)的建立代表了一次认证协议的顺利协商。它是一系列安全参数的集合，标识了某个密钥材料及其相关操作信息。MSA架构定义了PMK-MKD SA、PMK-MA SA和Key Holder Communication SA等安全关联来组织其密钥材料。

8.4.2 密钥体系

密钥体系是整个MSA安全架构的核心和最终目标。在一个应用了MSA架构的Mesh网络中，一个MP只有通过身份认证后建立起一套密钥体系才被允许在网络中参与通信。相比WLAN网络，Mesh网络的特性对密钥的保护提出了不同的要求。例如在自组织网络中作为认证中心节点的认证服务器并非随时可达，或者密钥材料的分发需要经过不可靠的多跳路径，这就要求密钥体系作相应调整。

如图8-5所示，MSA密钥体系结构可分为链路安全(Link Security)和密钥分发(Key Distribution)两个分支。Link Security分支用于MKD，MA和Supplicant MP间的密钥生成和分发，用于保护任以相邻MP对间的每跳数据链路的安全；Key Distribution分支用于保障Key Holder间的端到端安全通信以及PMK-MA密钥的安全分发。在结构上可以把Key Distribution分支想象成构建在Link Security分支上的一条安全隧道。

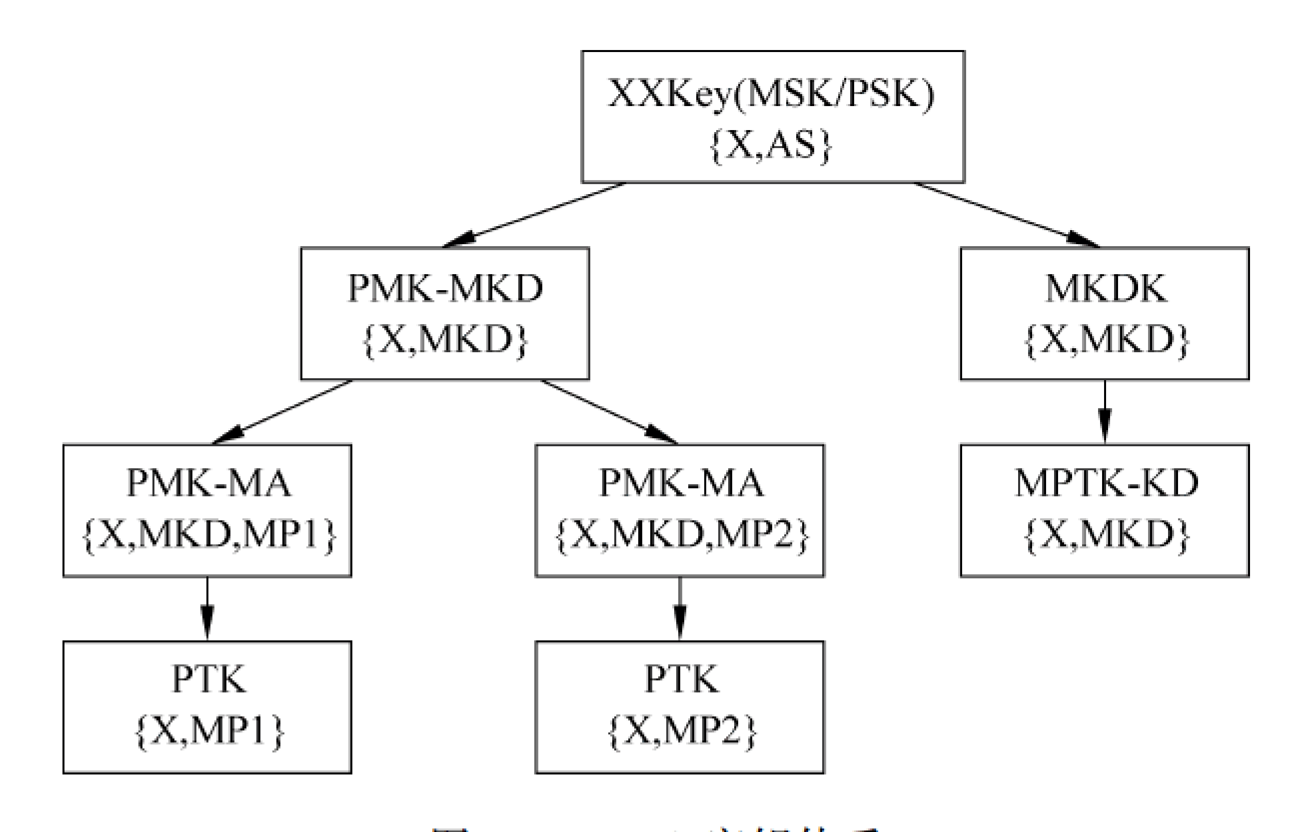


图8-5 MSA密钥体系

1. 链路安全分支建立

建立密钥体系之前在Supplicant MP和AS/MKD间必须已经存在一个共享的MSK(Master Session Key)或PSK(Pre-Shared Key)[23]。

(1) Level 1：MKD 和 Supplicant MP 分别使用 MSK 生成 PMK-MKD；MKD 生成一个随机数 ANONCE[24] 和标识符 PMK-MKDName，并建立一个 PMK-MKD SA。

PMK-MKD=KDF-256(MSK,"MKDKeyDerivation",MeshIDlen ‖ MeshID ‖ MKDDID ‖ 0x00 ‖ SPA)
PMK-MKDName=TK-128(SHA-256("MKDKeyName" ‖ MeshIDlen ‖ MeshID ‖ MKDDID ‖ 0x00 ‖ SPA ‖ ANonce))

PMK-MKD 由 MKD 和 Supplicant MP 共同持有，用于导出下层密钥，不参与具体的安全操作。每对 Supplicant MP 和 MKD 都可以分别计算相应的 PMK-MKD，MKD 上存在多个对应于不同 Supplicant MP 的 PMK-MKD，它们通过 SPA 区别并用 PMK-MKD Name 唯一标识。

(2) Level 2：MKD 针对不同的 MA-Supplicant MP 对生成 PMK-MA 和标识符 PMK-MAName，通过 Mesh 密钥传输协议将 PMK-MA 以及它选择的 ANONCE 分发到指定 MA。并建立一个 PMK-MA SA。

PMK-MA=KDF-256(PMK-MKD,"MAKeyDerivation",PMK-MKDName ‖ MAID ‖ 0x00 ‖ SPA)
PMK-MAName=TK-128(SHA-256("MAKeyName" ‖ PMK-MKDName ‖ MAID ‖ 0x00 ‖ SPA))

PMK-MA 作用与 802.11i 中的 PMK 类似，由每对 Supplicant MP 和 MKD 共同持有，并由 MKD 通过安全协议发送给对应的 MA 用以完成 MSA 四次握手。每个 MA 上也存在多个 PMK-MA，每个 PMK-MA 对应于不同的 Supplicant MP，它们通过 PMK-MKD，PMK-MKDName 和 MAID 区别并用 PMK-MAName 唯一标识。

(3) Level 3：MA 与 Candidate MP 使用 PMK-MA 生成 PTK。

PTK=KDF-PTKlen(PMK-MA,"PTKKeyDerivation",SNonce ‖ ANonce ‖ MAID ‖ SPA ‖ PMK-MAName)
PTKName=TK-128(SHA-256(PMK-MAName ‖ "PTKName",SNonce ‖ ANonce ‖ MAID ‖ SPA))

生成 PTK 使用的方法与 802.11i 四次握手中使用的方法基本一致。与 802.11i 不同的是，生成 PTK 使用的 ANONCE 可以是由 MKD 获得的而不是自己产生的。Supplicant MP 在获得 ANONCE 后按照与前面 MKD 一致的步骤生成 PMK-MA，并进一步生成 PTK。

2. 密钥分发分支建立

建立密钥体系之前在 Aspirant MA 与 MKD 间必须已经存在一个共享的 MSK(Master Session Key)或 PSK(Pre-Shared Key)。

(1) Level 1：MKD 使用 MSK 或 PSK 生成 Mesh KDK。

MKDK=KDF-256(MSK,"MKDK",MeshIDlen ‖ MeshID ‖ MKDDID ‖ MAID ‖ ANonce)
MKDKName=TK-128(SHA-256("MKDKName" ‖ MeshIDlen ‖ MeshID ‖ MKDDID ‖ MAID ‖ ANonce))

与 PMK-MKD 类似，MKDK 由 MKD 和 Aspirant MA 共同持有，用于导出下层密钥，不参与具体的安全操作。

(2) Level 2：MKD 使用 MKDK 和 MKDKName 分别生成 MPTK-KD 和 MPTK-KDName。

MPTK-KD=KDF-256(MKDK,"MPTK-KD",MANonce ‖ MKDNonce ‖ MAID ‖ MKDID)

MPTK-KDName= TK-128(SHA-256(MKDKName ‖ "MPTK-KDName" ‖ MANonce ‖ MKDNonce ‖ MAID ‖ MKDID))

MPTK-KD 是 MKD 和 MA 间的专用 PTK,用于保护两者间的密钥发送等 Key Holder 通信的保密性和数据完整性。MKD 和 MA 分别提供并交换一个 256bit 伪随机数 MKD-Nonce 和 MA-Nonce,与 MKDK 和双方 MAC 地址一起生成 MPTK-KD 和 MPTK-KDName。

8.4.3 MSA 协议集

1. 初始 MSA 认证

初始 MSA 认证机制实现了 Candidate MP 加入一个 Mesh 网络时必要的身份认证和链路安全分支密钥体系生成。

如图 8-6 所示,一个初始 MSA 认证过程大致可分为三个阶段:由一对双向的 Peer Link Open/Confirm 消息构成的 PLM(Peer Link Management)协议交互阶段;可选的 EAP 身份认证阶段;MSA 四次握手阶段。

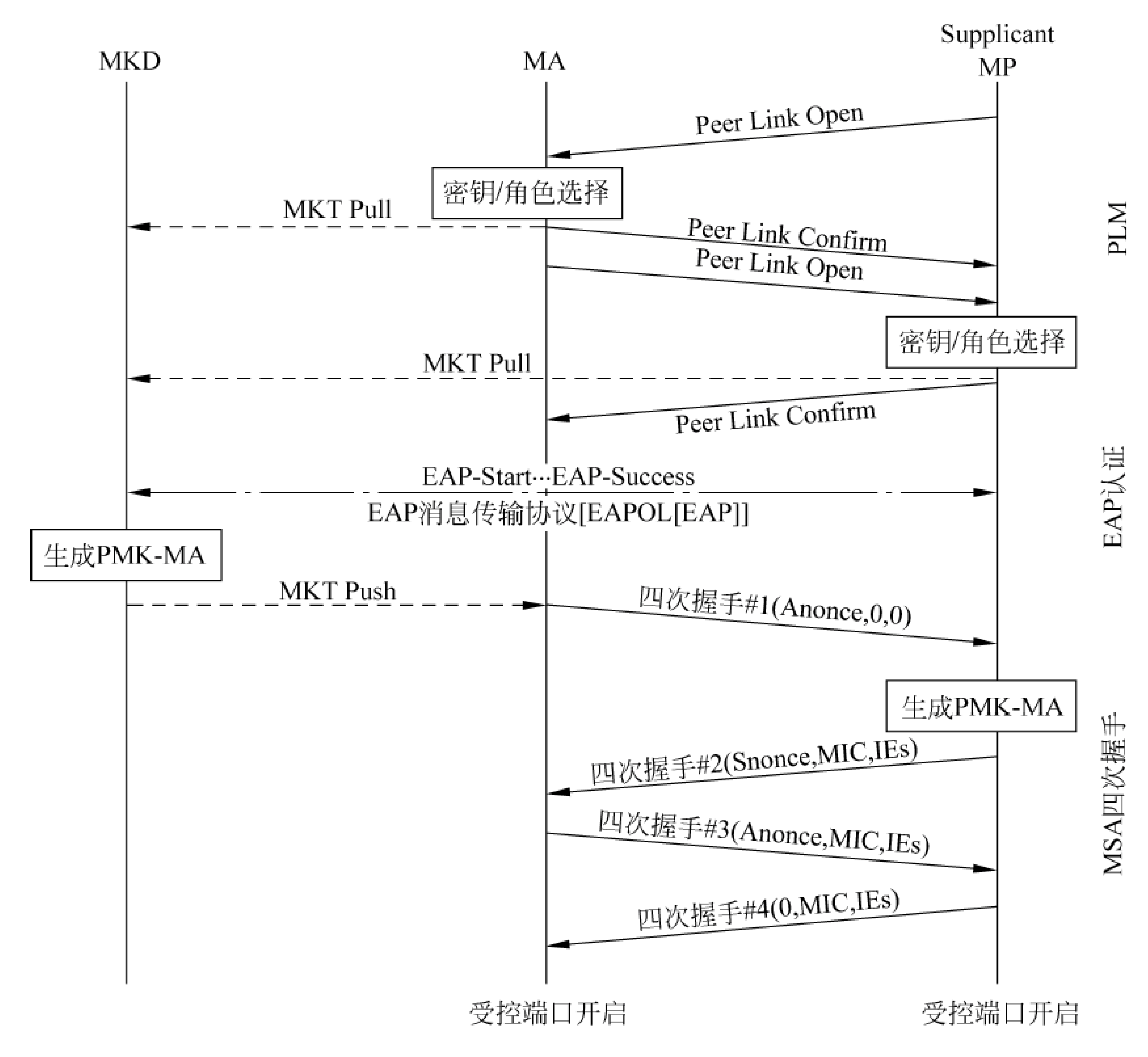


图 8-6 初始 MSA 认证过程

1) PLM 子协议

PLM 协议[25,26]类似与 802.11 关联阶段,在该阶段协商用于后续身份认证和密钥应用的一系列安全参数。另外由于 Mesh 网络的对等与自组织特性,在 PLM 阶段还将进行特殊的密钥选择和 EAP 角色选择操作。

（1）PLM有限状态机。PLM协议使用一个有限状态机来控制整个协议的运作。状态机包括7种状态。通过生成3种MLME事件（ACTOPN、PASOPN、CNCL）、5种消息处理事件（OPN_ACPT/RJCT、CNF_ACPT/RJCT、CLS_ACPT）或4种计时器事件（TOR1、TOR2、TOC、TOH）推动状态转移，完成发送各类消息帧和设置/清除计时器操作。由状态转移图可以看到，除了一般的消息收发控制，PLM状态机的重要功能即是消息乱序处理和超时重传控制。

由于Mesh网络的对等特性，PLM协议的通信双方可能以一种顺序的方式收发消息，也可能以一种同步的方式收发消息。这将导致消息的乱序，例如双方同时发出一条Open请求，或是Confirm消息先于Open消息到达。在设计PLM状态机时考虑到了所有可能情况，并通过设置适当的状态转移妥善地解决了这一问题。

PLM状态机定义了三个计时器来实现超时管理。一个Retry Timer（TOR1）用于控制Open消息的重传并使用可能的backoff算法避免冲突，当Open消息重传超过最大重传次数（TOR2）后，宣布PLM协商失败并进入HOLDING状态。一个Confirm Timer用于控制乱序状态下的等候时间，如果在收到Confirm消息后一定时间（TOC）没有完成Open消息的处理，则宣布PLM协商失败并进入HOLDING状态。一个Holding Timer用于控制在HOLDING状态的停留时间，在一定时间（TOH）内没有接到对方的Close应答则自动回到IDLE状态。

（2）消息帧处理[20]。由于自组织的Mesh网络中各节点可能无法事先得知对方的安全能力，例如是否有到MKD的连接、是否有可用的存储密钥等，因此简单的关联请求/关联响应两条消息无法为双方协商确定各自的角色。PLM对WLAN关联阶段进行补充使用一对异步Peer Link Open/Confirm消息构成以克服自组织网络中难以同步的问题。

在PLM交互阶段，各MP分别向对方发送Peer Link Open Action帧。Peer Link Open消息通过包含其中的各种信息元素（Information Element，IE）向对方通告各种后续协商所需的安全参数，这些IE例如通告支持对/组密钥/AKM集和本地存储密钥的RSN IE，通告MSA安全角色信息的MSC IE以及通告MSA能力、初始认证参数和选定对/组密钥、AKM集的MSA IE等。

在收到Peer Link Open消息后双方比较各自支持的密钥集并选择一个用于后续操作。如果从Peer Link Open消息中发现本地存储密钥的存在那么说明密钥体系已经建立，除非强制要求进行MSA初始认证否则可以使用Abbreviated握手机制。如果没有可用的存储密钥则双方进行802.1x角色协商，为后续的EAP认证作准备。

Peer Link Open消息处理无误后MP将向对方发送Peer Link Confirm消息确认选定的各种密钥集、是否进行初始MSA认证及角色协商结果。Peer Link Confirm消息主要用于保持同步和信息一致性，抵抗降级攻击，其中不含有新的信息。在对Peer Link Confirm消息进行验证并通过后，MP根据配置的策略进行EAP认证或直接加载密钥开始MSA四次握手。

（3）密钥选择。一对节点在之前的初始MSA认证担当的角色不同，则可能生成不同的PMK-MA，因此有可能通信双方在本地都保存了不止一个可用的PMK-MA。这样就需要进行一个密钥选择操作协商选定一个适合本次认证场景的存储密钥。如果密钥选择操作失败，那么说明必须进行初始MSA认证重新建立密钥体系。

密钥选择过程可以由一张 5 个输入一个输出的表[20]表示。

- valid-local-key：是否具有可用的本地存储密钥。
- cached-peer-key：是否具有可用的对方存储密钥。
- connected to MKD(peer)：对方是否维护着通往 MKD 的一条链路。
- connected to MKD(local)：本地是否维护着通往 MKD 的一条链路。
- selector：本地 MAC 地址是否高于对方 MAC 地址。

输出为一个本地或对方的一个 PMK-MA 标识符，如果不存在这样一个标识符则宣布协商失败。如果由输出标识符标识的密钥材料在本地无法找到，节点将调用密钥传输协议向 MKD 申请相关密钥。

(4) 角色选择。如果双方之前协商结果为需要进行 EAP 认证，则需要确定各自在 EAP 认证中的角色。角色选择过程也可以由一张 5 个输入的表[20]表示。

- require auth(peer)：本地是否要求强制认证。
- require auth(local)：对方是否要求强制认证。
- connected to MKD(peer)：对方是否维护着通往 MKD 的一条链路。
- connected to MKD(local)：本地是否维护着通往 MKD 的一条链路。
- selector：本地 MAC 地址是否高于对方 MAC 地址。

角色选择完成后则一个节点成为 Authenticator，而另一个成为 Supplicant，Mesh 节点由对等结构转变为暂时的 C/S 结构。

2) EAP 认证

如果需要进行 EAP 认证并且完成了角色选择操作，则通信双方将开始一个 802.1x/EAP 认证过程。在该过程中需要使用 EAP 消息传输协议封装原始的 EAPOL 数据帧，以适应多跳传输环境。认证成功后 Supplicant 和 MKD 分别计算 PMK-MKD；MKD 为 MA 生成 A Nonce 和 PMK-MA 并使用密钥传输协议将它们分发给 MA；最后 MKD 将为每个 MKD-Supplicant MP 对建立一个 PMK-MKD SA，并为每个 PMK-MA 建立一个 PMK-MA SA。

3) MSA 四次握手

MSA 四次握手与 802.11i 四次握手流程基本相同。不同的是 MA 发出的第一条消息中的 ANonce 是随 PMK-MA 一起从 MKD 获得的，Supplicant MP 在收到这个 ANonce 后才能够计算得到 PMK-MA。另外在第二、三条消息中还会附上 PLM 阶段 Open 消息中的几个信息元以确保信息的一致性，验证 PLM 各消息的完整性。

在四次握手结束后双方就完成了相互认证，建立并加载了密钥体系链路安全分支。双方打开 802.1x 控制端口，开始可靠的数据传输。

2. 简化握手协议

简化握手协议[27]用于在已经拥有存储 PMK-MA 的情况下建立经认证的 MP 间点到点链路及会话密钥。当一个 MP 认为自己与对方之间至少有一个共享的 PMK-MA 且对方支持简化握手协议时，该 MP 可以发起简化握手协议简化认证过程。通信双方在两种情况下共享 PMK-MA。

(1) 两个 MP 直接进行过相互认证。

(2) 其中一个 MP 由 MKD 获得另一个 MP 的 PMK-MA。

如图 8-7 所示，简化握手协议使用为 PLM 协议规定的 Action 帧（Open、Confirm）进行信息交互。简化握手将原有的 PLM 交互和四次握手合二为一，取代原来至少八条消息组成的 MSA 认证机制。简化握手功能主要通过其中携带的 RSN IE 和 MSA IE 的交换实现，在 Peer Link Open 帧中加入携带原来四次握手传递的 Nonce 值，PMK-MA 选择协商列表，安全能力协商信息以及加密的组密钥等信息的 IE，使通信双方在交换 Peer Link Open 消息后就可以使用共享 PMK-MA 生成 PTK 了。

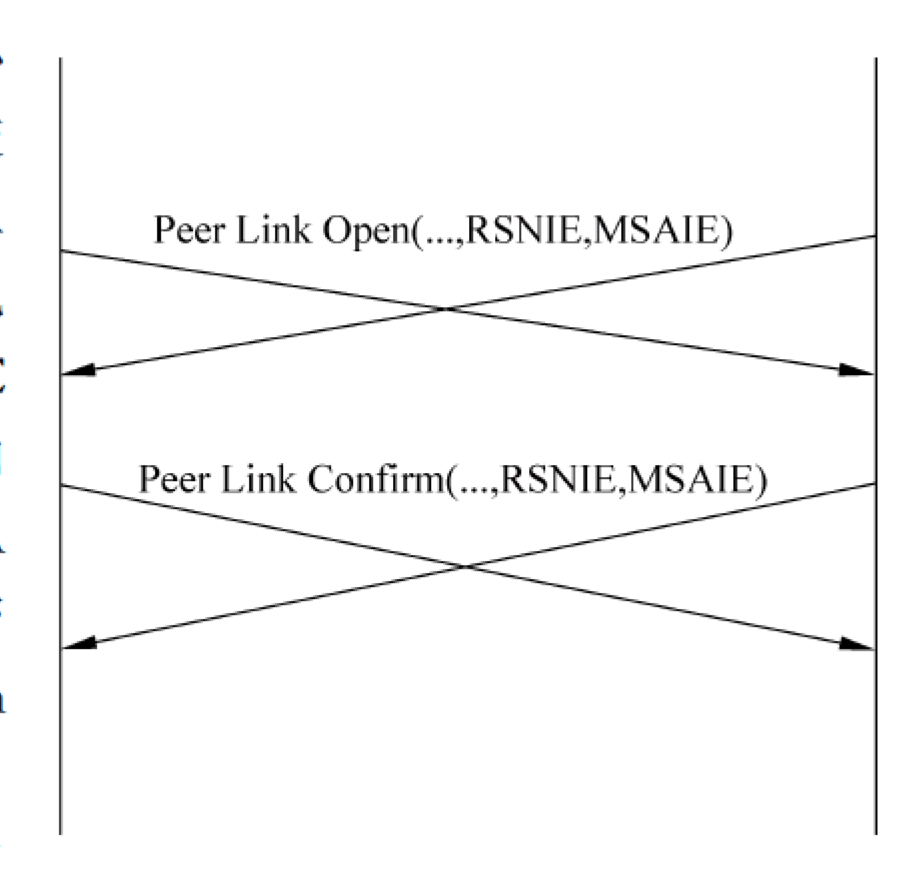


图 8-7　简化握手认证协议

简化握手功能主要包括 PMK 协商，安全能力协商和密钥管理：

- PMK 协商功能选择用于简化握手的 PMK-MA。如果 PMK 协商无法选择一个可用的 PMK-MA 则简化握手失败。如果 PMK 协商选定了一个不同于先前提议的 PMK-MA，两个 MP 将使用更新的 PMK-MA 参数进行开始新的简化握手实例，或执行初始 MSA 认证获得一个新的 PMK-MA SA。
- 安全能力协商功能用于建立安全关联的安全参数，例如成对密码零件（pairwise cipher suite）、AKM suite 和其他相关参数。
- 密钥管理功能生产 KEK、KCK 和 TK，并将各 MP 的 GTK 分发给对方。

简化握手可以使用与共享 PMK-MA 绑定的 KEK 和 KCK 对 Peer Link Open 帧中携带的密钥信息进行机密性和完整性保护，在不降低安全性的原则上最大限度地简化了协议，减小了消息个数及处理开销，提高了协议执行的效率。

3. Key Holder 安全关联握手协议

MP 可以通过 MKHSH 协议建立与 MKD 间的 Mesh Key Holder 安全关联，使自己成为 MA。这个安全关联代表了密钥体系的密钥分发分支，保证两者间的所有 MSA Key Holder 通信的安全。建立 KH 安全关联分为两个阶段：MKD 发现和 MP 发起的 MKHSH 协议。随着安全关联的建立将生成 MPTK-KD 保护所有通信报文和传输的密钥。该协议的前提是 MKD 与 MA 之间共享 MKDK。

MKHSH 协议是一个四次握手类型的协议，主要交换了用于计算 MPTK-KD 的 MA Nonce 和 MKD Nonce，以及协商双方支持的 Key Holder 传输类型。消息通过 MIC 机制使用 KCK-KD 和选定的 MIC 算法对消息帧中的各个域和 IE 进行完整性保护。整个 Mesh Key Holder Security Handshake 协议由四条消息[28]组成。

1）消息 1：MP→MKD

传输 MA-Nonce。收到有效的 Message 1 后 MKD 判断 MP 是否能授权成为 MA。如果通过则选择随机数 MKD-Nonce 会同 MA-Nonce 计算 MPTK-KD，最后发送消息 2。

2）消息 2：MKD→MP

传输 MKD Nonce、MKD 支持的 Key Holder 传输机制。消息受由 MKD 生成的 MIC 值保护。MP 在收到有效的消息 2 后就可以计算 MPTK-KD 并选择一个 Key Holder 传输类型了。

3）消息 3：MP→MKD

确认各种信息以保证信息一致性。如果握手出错则使用 Status Code 表明错误原因。消息受由 MP 生成的 MIC 值保护。

4）消息 4：MKD→MP

这条消息不参与认证或 MPTK-KD 生成，它仅仅为了使 MP 能够独自管理握手消息的重传，而 MKD 只负责对收到的消息进行回复。消息受由 MKD 生成的 MIC 值保护。

成功握手后双方均使用持有的 MA Nonce 和 MKD Nonce 计算出 MPTK-KD，并且协商确定使用的 Key Holder 传输类型。Key Holder 安全关联建立，MP 成为 MA 并可以为其他 Supplicant MP 提供认证服务。

4．Key Holder 安全关联解除协议

一个网络中可以存在多个 MKD 域，但是一个 MA 在某一时刻只能属于一个 MKD 域，也即只能与一个 MKD 建立 Key Holder 安全关联。这样当一个 MA 希望移动到另一个 MKD 域时除了建立新的 Key Holder 安全关联还需要删除原有的 Key Holder 安全关联。

MKHST 协议就用于 Mesh Key Holder 间协商删除已建立的安全关联。在诸如 MKD 不再提供服务的情况下它可以由 MKD 发起，而在 MA 需要与离开当前 MKD 域并与另一个 MKD 建立安全关联的情况下也可以由 MA 发起。这时 MA 需要先通过 MKHSH 与新的 MKD 建立安全关联，然后再发起 MKHST 删除与旧 MKD 间的安全关联。

MKHST 协议由一个请求和一个响应两条消息组成[29]。协议发起方称为 Teardown Requester，另一方称为 Teardown Responder。Requester 在请求消息中表明解除原因以及需要删除的安全关联，安全关联使用每条消息包含的 MPTK-KDShortName 进行标识，同时删除的还包括 MPTK-KD 和与之相关的 Replay Counter。两条消息均受到 MIC 保护。

5．Mesh 密钥传输协议

如前文所提到的，在初始 MSA 认证中 MA 可能会要求 MKD 提供某个 PMK-MA 密钥，MKD 也需要将生成的密钥分发到 MA。为此 MSA 中特别规定了一套 Mesh 密钥传输协议，应用密钥分发分支的 MPTK-KD 保护 MKD 到 MA 的 PMK-MA 传输过程以及密钥撤销过程。Mesh 密钥传输协议由三个子协议组成，分别用于几种特定的场景。

（1）Pull Protocol[28]：Pull 协议适用于一般场景。在密钥过期或对 Supplicant MP 进行认证时在本地找不到所需的 PMK-MA 情况下，MA 就发送一个 PMK-MA request 消息申请密钥。MKD 收到 PMK-MA request 后检查 MIC 和重放计数器，若无误则根据其中携带的 PMK-MAName 找到对应的 PMK-MA，并在 PMK-MA response 消息中返回 PMK-MA。若根据 PMK-MA Name 找不到合适的 PMK-MA，则 MKD 将标明相应 response 消息类型并把数据帧中密钥信息域置 0。这两条消息都受完整性保护，PMK-MA 被加密传输。

（2）Push Protocol：Push 协议主要用于拓扑建立阶段。首先由 MKD 向 MA 发送一条含有 PMK-MA 标识信息的 PMK-MA Notification 消息，而后 MA 发起完整 Pull 协议申请 Notification 中标识的 PMK-MA。三条消息都受完整性保护，PMK-MA 被加密传输。

（3）Key Delete Protocol：当 PMK-MA 过期需要撤销时，由 MKD 发送一条密钥撤销请求 PMK-MA delete。MA 收到 PMK-MA delete 消息后，将验证 MIC 以及重放计数器，

通过后 MA 使用 PMK-MA delete 消息中包含的 PMK-MKDName 和 SPA 计算 PMK-MAName，并且撤销由这个 PMK-MAName 所标识的 PMK-MA。最后向 MKD 发送类型为“Key delete acknowledged”的 PMK-MA response 消息确认密钥撤销。两条消息都受完整性保护。

6. EAP 消息传输协议

为了使 EAP 消息能够适应 Mesh 网络的多跳传输环境，避免将 MP 间中继的 EAP 消息和自发的 EAP 消息混淆，MSA 架构定义了 Mesh EAP 封装帧封装 EAPOL 消息，并使用 Mesh EAP 消息传输协议在 MKD 和 MA 之间传输。

MA 使用 EAP encapsulation request message[28] 向 MKD 发送来自 Supplicant 的 EAP 消息，或者向 MKD 请求发起 EAP 通信(EAP-Start)。MKD 收到 EAP encapsulation request message 后验证 MIC，储存重放计数器值用于生成回复，然后提取里面封装的 EAP 消息并转发给 AS。

当 MKD 由 AS 接到回复 EAP 消息后，使用 EAP encapsulation response message 封装后发送给 MA。该消息可以有三种类型 response、accept 或 reject，分别用于封装 EAP Response、EAP Accept 和 EAP Reject 数据。MA 收到 EAP encapsulation response message 后验证 MIC 以及收到的重放计数器值与最近发出的是否匹配，并根据最后一个 response 消息的类型判断需要进行的进一步操作。

8.4.4　安全方案协议协作实例

实际的安全方案是多个协议协作的过程，如图 8-8 所示[30]。Candidate MP 希望加入 WLAN Mesh 网络，就向其无线传输范围内的 MA 1 发起初始 MSA 认证。在 PLM 阶段通信双方协商确定使用的密钥算法集、没有可用的存储密钥，并且由 Candidate MP 作为 EAP 认证的 Supplicant 端。接着双方开始 EAP 交互，MA 1 使用 EAP 消息传输协议与 MKD 1 交换 EAP 消息，MKD 1 处理 EAP 消息并与 AS 使用 Radius 协议认证 Candidate MP，并使用密钥传输 Push 协议分发生成的 PMK-MA 至 MA 1。成功后 Candidate MP 和 MA 1 进入四次握手阶段生成 PTK 保护之间的数据传输，Candidate MP 加入网络。

MA 3 希望加入 MKD 1 域，它将先通过 Key Holder 安全握手与 MKD 1 建立安全关联，然后再发起 Key Holder 安全解除协议取消与 MKD 2 间的安全关联。

MAP 与 MA 2 之间曾经进行过初始 MSA 认证生成了共享 PMK-MA，因此当需要重新生成 PTK 时双方只需要使用简化握手协议。如果 MA 2 没有在本地找到 MAP 的密钥材料，它将通过密钥传输 Pull 协议向 MKD 1 请求 MAP 的 PMK-MA，再重新发起简化握手协议。

8.4.5　协议安全性分析

MSA 认证方案由 802.11i 方案发展而来，但是与 802.11i 方案有着本质的不同，主要体现在 Mesh 网络的对等特性上。这意味着网络中的节点在不同的交互中可能担任不同的安全角色，协议交互消息的顺序也不再是确定的了。这就要求人为地明确消息的方向性以阻止重放攻击。

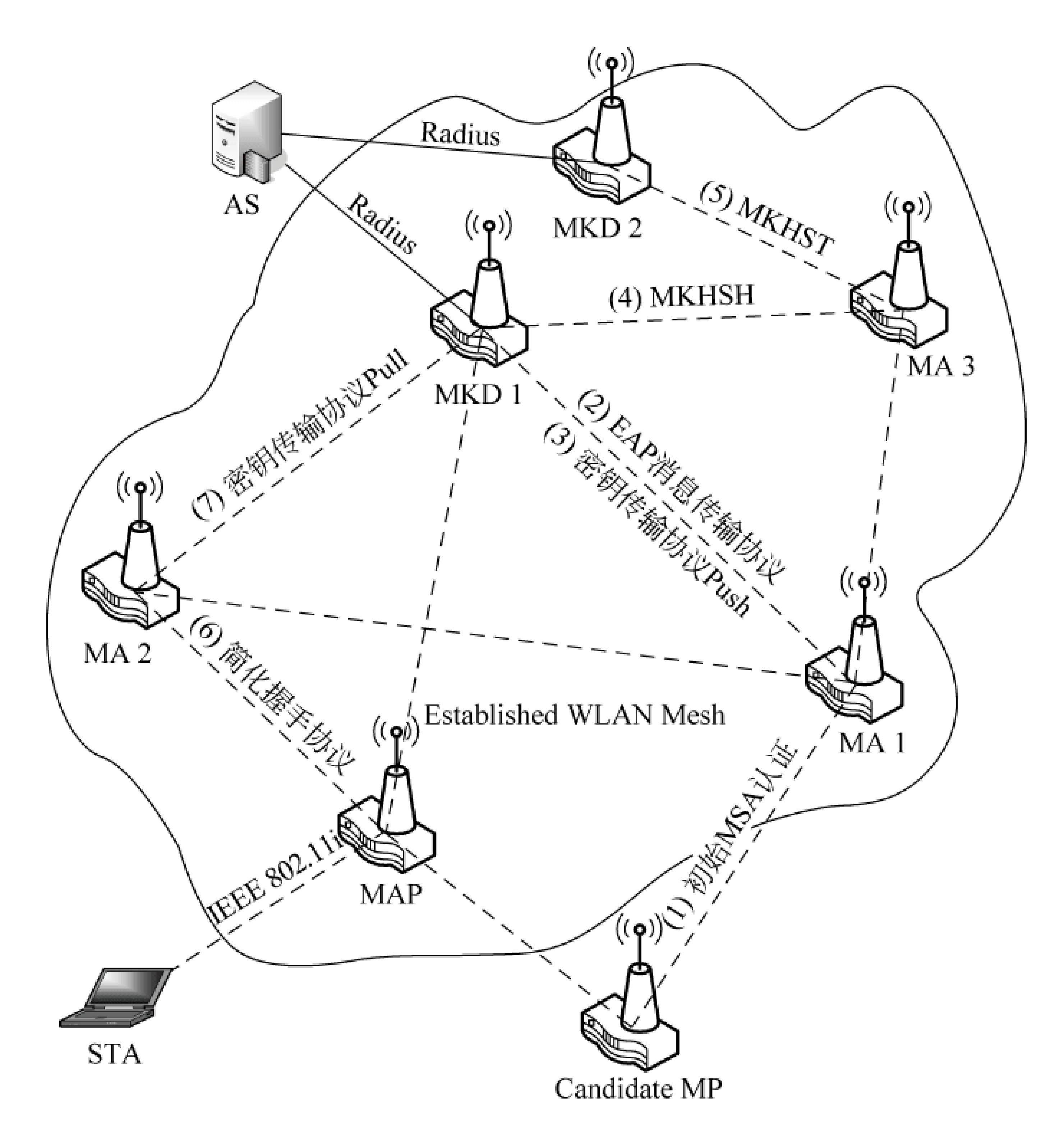


图 8-8　MSA 协议协作

初始 MSA 认证协议分为三个阶段。在 PLM 交互阶段由于密钥体系还未建立因此无法对其进行有效保护，但是该阶段交互的信息的完整性可以在四次握手阶段完成后得到验证。EAP 框架的安全性在文献[31]中得到证明。

MSA 四次握手和 802.11i 四次握手有着几个相同的安全目标：

(1) 通过正确的 MIC 编解码验证双方持有相同的 PMK-MA 并保护消息完整性；

(2) 通过 NONCE 的选择交换确保生成 PTK 的新鲜性；

(3) 使用冗余的第 4 条消息同步双方向 MAC 层加载 PTK/GTK 的操作。

与 802.11i 四次握手不同的是，MSA 四次握手除了需要在消息中添加 PLM 阶段保存的信息验证其真实性和完整性，还需要明确消息的方向性。通过规定消息中 NONCE 排列的顺序，MSA 架构在不添加新状态的情况下巧妙地解决了方向性的问题。

简化握手协议使用 Peer Link Open/Confirm 消息进行通信，但是由于存储 PMK-MA 已经选定，因此能够对消息携带的密钥信息进行加密，并对传输的消息使用 MIC 完整性保护。但是由于使用的 KEK 和 KCK 是和 PMK-MA 绑定的，无法保证密钥新鲜性。

Key Holder 安全握手同属于四次握手类型协议，与 MSA 四次握手具有类似的安全目标和手段。然而 MSA 草案专门规定了消息中携带的 MA 和 MKD 地址排列顺序，更直接地明确消息的发起者和接收者，以防止重放攻击。

Key Holder 安全解除协议以及密钥传输协议中的三个子协议 Pull、Push 和 Delete 协议本质上都是请求/响应式协议，由于 Push 协议中的第一条消息是无需保护的。这些协议

的消息在应用了重放计数器的同时，明确加入了发起者或接收者地址，以避免回射攻击威胁。协议中的消息均使用 MIC 验证码，保护消息完整性。

文献[32]使用 SafeNet 和 PCL 技术对整个 MSA 的密钥体系进行了详细的安全分析，可以证明 MSA 密钥体系在现有单 MKD 架构 Mesh 网络中是安全且足够有效的。

8.5 小结

本章首先对无线 Mesh 网络进行了概述，其次对无线 Mesh 网络的安全风险和安全需求进行了分析，然后重点阐述了基于 MSA 协议的安全协议及相关技术。解释了 MSA 的一些基本概念，包括体系中定义的各种角色、安全关联，以及协议集概述；详细介绍 MSA 架构的核心：MSA 密钥体系。包括其各分支和层次结构，各种密钥的生成、用途以及应用场景；接着是 MSA 协议集的详细介绍，协议集中包括六种协议，用于完成用户身份认证及 MSA 密钥体系的生成、分发和维护。并希望通过一个协议协作实例加深读者的理解。最后对 MSA 协议进行了安全分析，得出 MSA 安全架构在现有单 MKD 架构 Mesh 网络中是安全且足够有效的结论。

参考文献

[1] 方旭明. 下一代无线因特网技术：无线 Mesh 网络. 北京：人民邮电出版社，2006.
[2] Akyildiz I F，Wang X，Wang W. Wireless Mesh Networks：a survey. Computer Networks. 2005 47(4)：445-487.
[3] Chlamtac I，Conti M，Liu J. Mobile Ad hoc Networking：Imperatives and Challenges. Ad hoc Networks. 2003 1：13-64.
[4] Draft for IEEE 802.11 ESS Mesh. IEEE P802.11s/D1.01.2007.
[5] Hauser J. Draft PAR for IEEE 802.11 ESS Mesh. IEEE Document Number：IEEE 802.11-03/759r2.
[6] 吴越，孙东来，易平等. 无线 Mesh 网络(SJTU-MESH)试验床的构建与应用. 电信科学，2008(9).
[7] Gu X，Hunt R，Wireless LAN Attacks and Vulnerabilities，In the Proceeding of IASTED Networks and Communication Systems，18-20 April，2005.
[8] Xia H，et al. Detecting and Blocking Unauthorized Access in Wi-Fi Networks.
[9] Parno B，Perrig A，Gligor V. Distributed Detection of Node Replication Attacks in Sensor Networks. Proc. IEEE Symp. Sec. and Privacy，2005.
[10] IEEE 802.11-04-0968r11. Issues for Mesh Media Access Coordination Component in 11s. Jan. 2005.
[11] Salem N B，Hubaux J P. Securing Wireless Mesh Networks. In IEEE Wireless Communication. April 2006，13(2)：50-55.
[12] Tchepnda C，Riguidel M. Distributed Trust Infrastructure and Trust-Security Articulation：Application to Heterogeneous Networks，in proceeding of 20th Int. Conference on Advanced Information Networking and Applications，(AINA)，April 2006，2：33-38.
[13] Garcia-Luna-Aceves J J，Mosko M. Multipath Routing in Wireless Mesh Networks. In first IEEE Workshop on Wireless Mesh Networks(WiMesh 2005). September 2005，Santa Clara，CA.
[14] Raya M，Hubaux J P. The Security of Vehicular Ad hoc Networks. Proc. SASN，2005.
[15] Salem N B，Hubaux J P. A Fair Scheduling for Wireless Mesh Networks. In Proceedings of WiMesh，September 2005，Santa Clara，CA.

[16] Gambiroza V, Sadeghi B, Knightly E. End-to-End Performance and Fairness in Multihop Wireless Backhaul Networks. Proc. MobiCom, 2004.

[17] Eastlake D 3rd. IEEE P802. 11s Call for Proposals. IEEE 802. 11s document, DCN 802. 11-04/1430r12, January 2005, http://www.802wirelessworld.com/index.jsp.

[18] Emeott S, Braskich T. Overview of Key Holder Security Association Teardown Mechanism. IEEE Document Number: IEEE 802. 11-07/2376r0 2007.

[19] Braskich T, Emeott S. Clarification and update of MSA overview and MKD functionality text. IEEE Document Number: IEEE 802. 11-07/2119r1 2007.

[20] Braskich T, Emeott S. Initial MSA comment resolution. IEEE Document Number: IEEE 802. 11-07/0564r2 2007.

[21] Zhao M, Walker J. Conner W S. Abbreviated Handshake Protocol Requirements. IEEE Document Number: IEEE 802. 11-07/0733r0 2007.

[22] Braskich T, Emeott S. Overview of Improvements to Key Holder Protocols. IEEE Document Number: IEEE 802. 11-07/1988r1 2007.

[23] Braskich T, Emeott S. Mesh Pre-Shared Key Clarification. IEEE Document Number: IEEE 802. 11-07/2037r0 2007.

[24] Braskich T, Emeott S. Key Hierarchy Nonce Update. IEEE Document Number: IEEE 802. 11-07/2649r0 2007.

[25] Zhao M, Walker J, Conner W S. Updates on Peer Link Management Protocol. IEEE Document Number: IEEE 802. 11-07/2174r0 2007.

[26] Zhao M. Resolutions for Comments on Minor Text Changes for Peer Link Management. IEEE Document Number: IEEE 802. 11-08/1082r3 2007.

[27] Zhao M, Walker J, Conner W S. Abbreviated Handshake for Authenticated Peer Link Establishment. IEEE Document Number: IEEE 802. 11-07/1999r3 2007.

[28] Braskich T, Emeott S. Mesh Key Holder Protocol Improvements. IEEE Document Number: IEEE 802. 11-07/1987r1 2007.

[29] Emeott S, Braskich T. Key Holder Security Association Teardown Mechanism. IEEE Document Number: IEEE 802. 11-07/2372r0 2007.

[30] 朱近丹. 基于 WLAN 的无线网状网络 WMN 安全接入协议研究. [硕士学位论文]上海交通大学, 2009. 1.

[31] He C, Sundararajan M, Datta A, et al. A Modular Correctness Proof of IEEE 802. 11i and TLS. In Proceedings of the 12th ACM Conference on Computer and Communications Security. ACM: Alexandria, 2005.

[32] Kuhlman D, Moriarty R, Braskich T, et al. In a Correctness Proof of a Mesh Security Architecture, Computer Security Foundations Symposium. 2008. CSF'08. IEEE 21st, 2008: 315-330.

第9章　对等网络及研究进展

CHAPTER 9

摘要：对等网络系统是一个新兴的研究领域，近些年得到迅速发展。本章首先对对等网络进行了概述性介绍，然后重点介绍了对等网络在路由、拓扑和查询这三方面的研究工作和研究进展。

关键字：P2P、DHT、路由、拓扑、查询。

9.1　P2P 概述

过去十年见证了 Internet 涌现出的大量分布式应用，其中最流行的应用之一是使用 P2P(Peer-to-Peer)系统的文件共享。在 1999 年 1 月，Shawn Fanning 离开 NorthWestern University，开发了软件 Napster[1]。Napster 是公众可用的第一个用于音乐共享的系统，并且取得了极大的成功。在它诞生不久，世界上就有几百万的用户使用它。然而，好景不长，美国唱片工业协会(RIAA)发起了对 Napster 的起诉，认为它在非法共享 MP3 音乐文件方面起了推波助澜的作用，因此违反了版权法。尽管与德国媒体公司 Bertlesmann AG 联合开发一个成员基于的分布系统以保障艺术家的版权收入，Napster 还是被司法局在 2001 年 3 月关闭。但 Napster 的结束并不意味着 P2P 文件共享的结束。相反，以 Gnutella[2] 为主力的分布式 P2P 系统仍然在继续发展并不断壮大。新系统如 KaZaA[3]、LimeWire[4]、Morpheus[5] 不断出现，P2P 用户数量持续快速增长。在 2000 年夏，KaZaA 软件已被超过 1 亿用户下载。此外，系统不再局限于共享音频、视频、软件和其他格式文件。它已经超越了文件共享的王国，出现了如 SETI@Home[36] 等利用系统参与者的空闲处理能力等新型应用。今天，P2P 系统已经在数据存储[6~8]、组通信[9,10]、消息系统[11] 等多方面得到了应用。

9.1.1　P2P 定义

P2P 正处于不断发展的阶段，因而并不存在一个精确的定义。当前给出的许多定义都是试图反映 P2P 系统发展过程中的某阶段的新特征。下面是有关文献给出的一些 P2P 定义。

Clay Shirkey：P2P 是一种利用位于 Internet 边缘的各种可用资源(如存

储空间、计算能力、媒体内容)的应用。访问这些分散的资源,就意味着要在连接不稳定和IP地址不可预见的环境里工作。由于网络上大量的节点工作在DNS系统之外,这些分散的资源具有不稳定的连通性和未知的IP地址。因此,P2P节点必须能够独立于DNS系统且高度自治[12]。

Mike Miler:P2P是一个网络体系,其中每个计算机有同等能力和责任。Miler定义了五个关键特性[13]:

(1) 网络提供节点间实时的数据传输或者消息传递。

(2) 节点即是客户端又是服务器。

(3) 网络的内容是由分布的节点提供。

(4) 节点具有网络控制权和自治权。

(5) 网络允许不总是连接的节点和可能没有永久IP地址的节点参与。

P2P工作组:P2P是通过在系统之间直接交换来共享计算机资源和服务。这些资源和服务包括信息交换、高速缓存、处理能力、存储空间。P2P可以整合这些经济的PC上计算机和网络连接,从而提供企业级的计算平台[14]。

一般地说,P2P是一个用于资源共享的peer群体,其中每个peer向群体提供资源同时作为回报从中获取所需资源。它的思想是基于世界上的事物是广泛分布且相互联系的,不可能通过一种集中化的方式管理如此庞大的结构。P2P通过分布于世界各地的个人计算机管理大量的计算能力、存储空间和连接。P2P中的每个peer自治又彼此依赖,所谓自治是指每个peer独立决定自己的行为而不受其他例如集中式授权机构的控制,同时每个peer又需要相互协作获得信息资源、计算资源,在本书中把每个peer称为节点,并把peer所组成的网络称为叠加网络(overlay network)。

9.1.2 P2P系统的分类

P2P系统具有多种分类标准。我们仅选取分散度和网络结构[15,16]作为分类标准,因为我们认为它恰当地反映了P2P系统的本质特征的归类。

分散度表示P2P系统中节点在进行相关信息搜索时依赖目录服务器的程度。它有三种情况:完全分散、部分分散、混合分散。完全分散的情况下,所有节点都是平等的,没有任何一个会比其他一个更重要些,不存在目录服务器;部分分散的情况下,一些所谓的超级节点充当了部分目录服务的功能以及改善系统性能;混合分散的情况下,整个系统依赖于一个或者非常少的不可替代的节点提供集中式的目录服务,而系统内其余的节点,具有彼此等同的功能。

网络结构是从拓扑透视的角度来看系统,它表示P2P系统的覆盖网络结构是受某种机制约束还是完全动态、即时的。前者为结构化,后者为非结构化。非结构化P2P系统中对等节点连接任意其他对等节点构成对等网络,数据放置与网络拓扑无关,系统中每一个节点只负责管理自己的数据文件,从而它的网络拓扑是一个随机连接图,不能保证一定可以找到目标节点。结构化P2P系统的网络有某种预定的结构如环或者网格,系统中每一个数据文件所放置位置由特定的协议确定,并提供了文件标识ID和文件存储位置之间的映射关系,从而保证具有有效路由,并能够保证最终找到目标节点。

结构化系统采用分布式哈希表(Distributed Hash Table,DHT)技术构造,其本质是完全分散构造。因此,本书仅对非结构化系统区分分散度情况。当前的系统中 Chord、CAN、Pastry 和 Tapestry 是结构化 P2P 系统。而非结构化根据分散度区分为以下三种系统:

(1) Napster 是混合分散的非结构化 P2P 系统。

(2) Gnutella、Freenet 是完全分散的非结构化 P2P 系统。

(3) Kazza 和 Morpheus 是部分分散的非结构化 P2P 系统。根据此标准,P2P 系统的划分如图 9-1 所示。

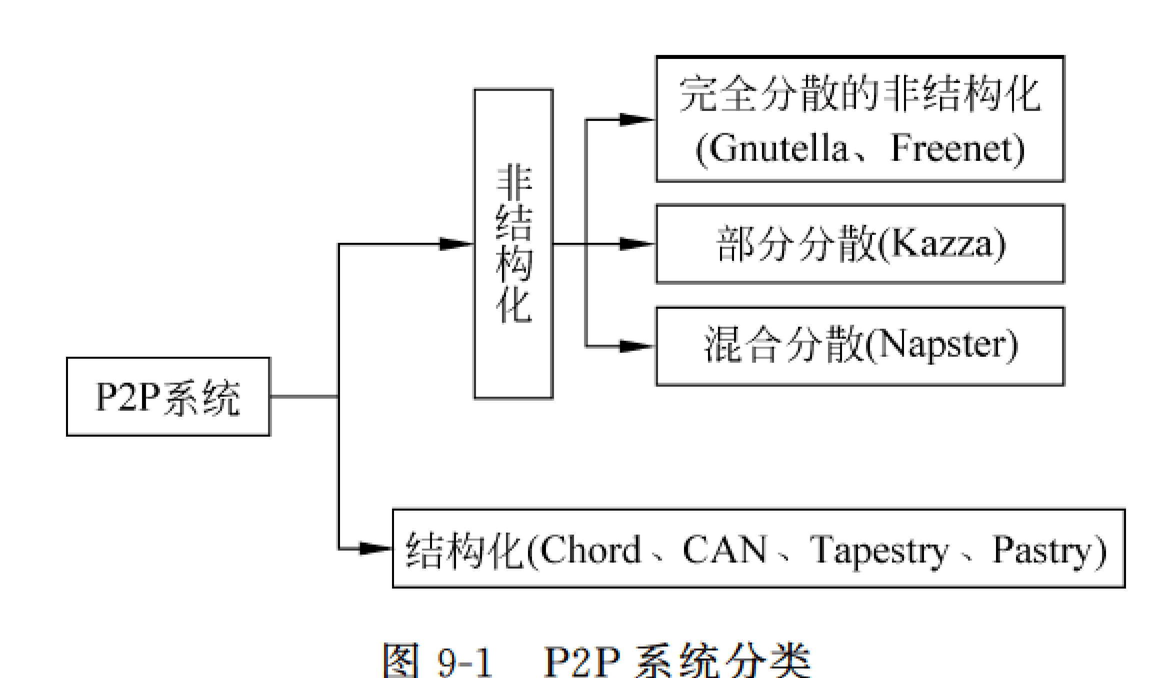


图 9-1　P2P 系统分类

9.1.3　P2P 系统的发展

在网络和分布领域中,术语 P2P 是一个相对新的名词。自从 1999 年 Napster 出现及 2000 年 P2P 成为研究热点[17],P2P 系统得到迅速发展。本节按时间出现的先后顺序,将 P2P 系统的发展大致分为二代。每一代都有其显著的特征:第一代网络结构为非结构化,第二代网络结构为结构化;第一代特点是采用泛洪的不确定性查询,第二代特点是采用基于分布式哈希表的确定性查询。在此,我们仅仅给出一个轮廓般的介绍,以刻画出其发展的轨迹,详细介绍请参考第 2 章。

1. 第一代 P2P 系统

第一代 P2P 系统是以非结构化为特征。它是伴随文件共享应用 Napster[1] 而开始的。Napster 的主要贡献是介绍了一种新型的网络体系,这种体系不同于传统的客户端/服务器架构,而是让计算机间借助一中心服务器可以自由通信交流。在某种程度上,计算机既是资源的消费者又是资源的提供者,因而术语 peer 更适合表示系统内的参与者的对等情况。为了在共享空间定位文件,Napster 的解决途径是提供了一个中心目录服务器。因而,Napster 系统由两种服务组成:存储服务和目录服务。存储服务是分布的,采用了 P2P 风格,而目录服务是集中式的,因此它是一种混合分散构造。但 Napster 仅是短暂存在,它的体系结构中集中式的目录服务是个关键性问题。首先,它存在单点故障问题;其次,尽管目录服务提供了定位的低开销,但是目录服务器上的负载也是随参与者数目线性增长,从而使得它是不可扩展的。

Napster 集中目录服务器瓶颈导致 P2P 新系统设计关注于如何减少这种集中目录服务,即趋向更高的分散度。Gnutella[2] 和 Freenet[18] 是新的系统代表,它们是完全分散的非结构化系统。这些系统中新的参与者必须知道一个已经参与系统的成员,然后使用一种泛

洪算法去得到关于别的参与者的情况，以建立邻居关系表。而对于一个给定的查询，也是采用泛洪算法。查询请求节点先将此请求泛洪到它的所有邻居，然后它的邻居再做类似的泛洪操作，直至查询发现。为了防止泛洪过度，请求包附加上 TTL 限制值。当此 TTL 值为 0 时，查询立即终止。

新的系统解决了集中瓶颈问题。然而，可扩展问题却更为严重。这是由于泛洪算法带来了高的网络流量，此研究可发现在[19,20]。而且，由于搜索范围受到限制，不能够确定性的发现存在 Gnutella 网络中的一个数据项或者一个资源文件。Freenet 采用了一个略好些的方案，它是基于文档路由模型。它利用哈希技术给数据项唯一标识 ID。当数据项插入系统时是尽可以插入到具有与它的标识最接近的节点。查询是由数据项的标识所引导定位的。但由于 Freenet 网络的随机特性，发现文档的概率并不高。

一些系统在混合分散和完全分散做了折中，采用了部分分散机制，如 KaZaA[3]。KaZaA 采用超级节点(super peer)技术，允许一些节点充当目录服务从而减少了需要定位数据的泛洪消息数。尽管如此，查询仍然是不确定的泛洪方式。

2. 第二代 P2P 系统

第二代的 P2P 系统是以结构化为特征，它是由 Chord[21]、CAN[22]、Pastry[23]、Tapestry[24] 等开始的，系统是完全分散组织。在这些系统中，一个共有的关键特性是基于分布式哈希表机制。在这些系统中，节点通过它的一些唯一性属性如 IP 地址哈希得到唯一标识 ID。数据项(data item)是以键值对＜key，value＞的方式表示，其中键是对于数据项的索引，而值可以是数据项的定位地址如 IP 或者 URL。通过哈希赋予数据索引键以唯一标识，并将此键对应的键值插入与此键标识最邻近的节点。注意，节点和键的哈希空间是相同的。查询时，通过将查询的键哈希得到唯一标识，并通过此唯一标识找到与之最邻近的节点(此节点存储了数据项所在的地址)。为了支持基于标识(ID)的查找，节点将哈希得到的 ID 空间组织成一个结构化的拓扑如 Chord 中的环(circle)、CAN 中的超环(torus)、Tapestry 的树(tree)。由于这种结构化拓扑，使得数据查找具有可确定性，即数据只要存在叠加网络上就可以以高概率确定性查找到。

当前两代 P2P 系统并存且都得到相应的发展。它们适应于不同的方面，互为补充。第一代系统在文件共享等并不需要确定性的查询结果的领域会有更广阔的前途。因为非结构化性提供松散的组织特性，可以让参与者有充分自由度，在最大限度减少对 P2P 系统的影响。第二代系统极大拓展了 P2P 系统的应用领域，使 P2P 系统成为分布计算的一个良好的平台，甚至成为下一代 Internet 应用的基础[25]。这是由于采用分布式哈希表在完全分布的环境下具有高度确定性和高度容错的特性，因而它能够更好的适应大多数分布式应用的要求。

9.2 分布式哈希表与 P2P 系统

9.2.1 分布式哈希表简史和技术原理

分布式哈希表是作为可扩展的分布式数据结构(Scalable Distributed Data Structure，SDDS)在 20 世纪 90 年代得到广泛研究的，其中术语 SDDS 是在 Litwin 等的经典论文[26]提

出的。但是，这些分布哈希表有中心部件，设计面向的是小规模的集群。面向大规模集群的分布式哈希表由 Gribble 等[27]采用 Java 实现，具有高度可扩展、容错和可用性。

分布式哈希表技术应用到对等网络是在 2001 年。一系列的基于分布式哈希表的 P2P 系统[21~24]被提出。这些系统采用分布式哈希技术能够支持上百万的机器动态参与，具有高度的可扩展性和容错性。目前，基于分布式哈希表技术构造的 P2P 系统已经成为 P2P 研究的一个重要内容。

下面简单介绍一下 P2P 系统中分布式哈希表技术原理。

哈希表是计算机科学里常见的数据结构，它能够根据索引的关键码值快速查找记录。它通常提供了两种基本功能：put 和 get 操作，也就是 put(key，value)，value＝get(key)，且平均查找时间为常量度 $O(1)$。分布式哈希表是哈希表的分布式构造，将哈希表由单个节点扩展到 Internet 上。由于分布式哈希表是布置在整个 Internet 上，put 和 get 操作需要借助于分布路由而实现。哈希表与分布哈希表的对照如表 9-1 所示。

表 9-1　哈希表与分布哈希表对照

哈　希　表	分布式哈希表
Key＝hash(data)	Key＝hash(data)
Put(key，value)	Lookup(key)－＞node_IP
Get(key)－＞value	Route(node_IP，put，key，value)
	Route(node_IP，get，key)－＞value

在 P2P 系统中是怎样采用分布式哈希表实现资源定位呢？

具体方法是首先将网络中的每一个节点分配虚拟地址标识(nodeId)，同时用一个关键字 Key 来表示其可提供的共享内容。取一个哈希函数，这个函数可以将 Key 置换成一个哈希值 H(key)。网络中节点相邻的定义是哈希值相邻。发布信息的时候就把(key，nodeId)二元组发布具有和 H(key)相近地址的节点上去，其中 nodeId 指出了文档的存储位置。资源定位的时候，就可以根据 H(key)相近的节点快速获取二元组(key，nodeId)，从而获得文档的存储位置。不同的 DHT 算法决定了 P2P 网络的逻辑拓扑，比如 CAN 是一个超环，而 Chord 是一个环，Tapestry 是树状。

分布式哈希表提供给 P2P 系统设计以非常简洁高效的接口，然而，在实际应用中，分布哈希表的设计要考虑如下三个关键问题：

- 唯一性：哈希函数必须避开碰撞或者碰撞概率非常小以至忽略不计。
- 动态性：哈希函数必须能够支持节点动态加入和离开的一致性。
- 合理尺寸：哈希表必须能够支持可扩展性，不能在单个节点上占用太大的空间。

对于第一个问题，可以通过定义一个固定的足够大的哈希空间，从而所有哈希值落在此空间而不依赖节点的数目，从而在静态或者动态下能够避开碰撞。对于第二个问题通常采用的是一致性哈希方案。一致性哈希最初是由 MIT 的 Karger 等[28]引入以解决分布 Cache 中的热点问题，目前已经被扩展到很多领域，特别是 P2P 计算中。一致性哈希是哈希中的一种，但它具有如下特性以适应动态分布的应用：

- 可扩展性。

• 负载均衡。
• 平滑性。

这些特性可以很好的适应动态性下的哈希正确性和一致性。对于第三个问题，解决的方法是将整个哈希表均匀分布到各节点上，每一节点负责它在哈希空间标识邻近的键值。因而需要结构化的拓扑组织以支持哈希表的分布及路由。故分布式哈希表构造是一种结构化的拓扑构造。

9.2.2 基于分布式哈希表的 P2P 系统/DHT-P2P 系统

DHT 技术对于 P2P 系统设计有革命性的影响。由于 DHT 支持具有有序的、高可扩展、可确定性查找的结构化拓扑，超越了之前的随机、Ad hoc 的非结构拓扑，涌现了一大批的新型 P2P 系统如 CAN[22]、Chord[21]、Tapestry[24]、Pastry[23]、Kademlia[29]、Symphony[30]、Viceroy[31]、Koorde[32]、P-Grid[33]等。这些系统都是采用 DHT 技术作为其设计基础。由于节点和数据都是通过分布式哈希组织成一个叠加网络，从而这些系统被称为基于 DHT 的 P2P 系统，一些文献也简称为 DHTs 或者 DHT。为了明确和简化，本文称之为 DHT-P2P 系统，在不至于混淆情况下，本文中也称之为 DHT 系统。

如上节所述，DHT-P2P 系统中每个节点负责一定范围的键值。叠加网中的节点既能够存储资源本身，也可能仅存储资源的地址指针，取决于算法需要。DHT-P2P 系统的路由是通过贪婪算法逐步逼近具有在 metric 空间最近距离的目标节点，这个 metric 空间取决于系统的拓扑组织结构。

9.2.3 DHT-P2P 系统特性

DHT-P2P 系统是结构化系统，相比较非结构化系统，它提供了几个很好的特性：

(1) 搜索不需要依赖于泛洪机制，因此造成较小的网络流量，大部分 DHT-P2P 系统中每个查询都只是需要 $O(\log n)$个消息和跳数。

(2) 每个查找请求都能以很高的概率解析，并且所需要的资源消耗是可预测的，而在非结构化系统里如果所请求的文档超越了查找所能覆盖的范围则查询失败，而且即使查找成功其资源消耗也不可预测。

(3) 搜索结果是确定性的。一方面结构化拓扑数据只要存在叠加网络上就可以以高概率确定性查找到，而非结构化受 TTL 限制易查找失败；另一方面查找结果也有确定性，而在非结构化系统中，不同的节点提交同样的搜索请求时很可能获得的结果也不同。

由于 DHT-P2P 系统这些的特性，使得它特别适合 Internet 分布应用，引起学术界的高度重视。

9.3 P2P 系统的典型代表

本节对每一代的典型系统进行介绍，以刻画出当代 P2P 系统的特征。

9.3.1　第一代 P2P 系统

1. Napster

Napster 是第一个公众使用的 P2P 系统。它采用一个中心服务器维护一个(文件名,IP 地址的)索引对,并且保持跟踪系统里所有音乐文件的定位,如图 9-2(a)所示。当一个用户想要一个特定的音乐文件时,它查询服务器得到存放这个音乐文件的机器 IP 地址,并从这台机器那里下载此音乐文件。文件的传输是分布式并且独立于中心索引服务器。这种简单途径的一个潜在的问题是缺乏可扩展性。然而,如果能够采用大规模服务器集群,如 Google 和 Yahoo,则可以达到相当大的可扩展性。尽管如此,由于集群系统需要相当大的投资在高性能机器和高的带宽等,因而并不适用于 P2P 网络。此外,这种途径也不是容错的,存在单点故障。

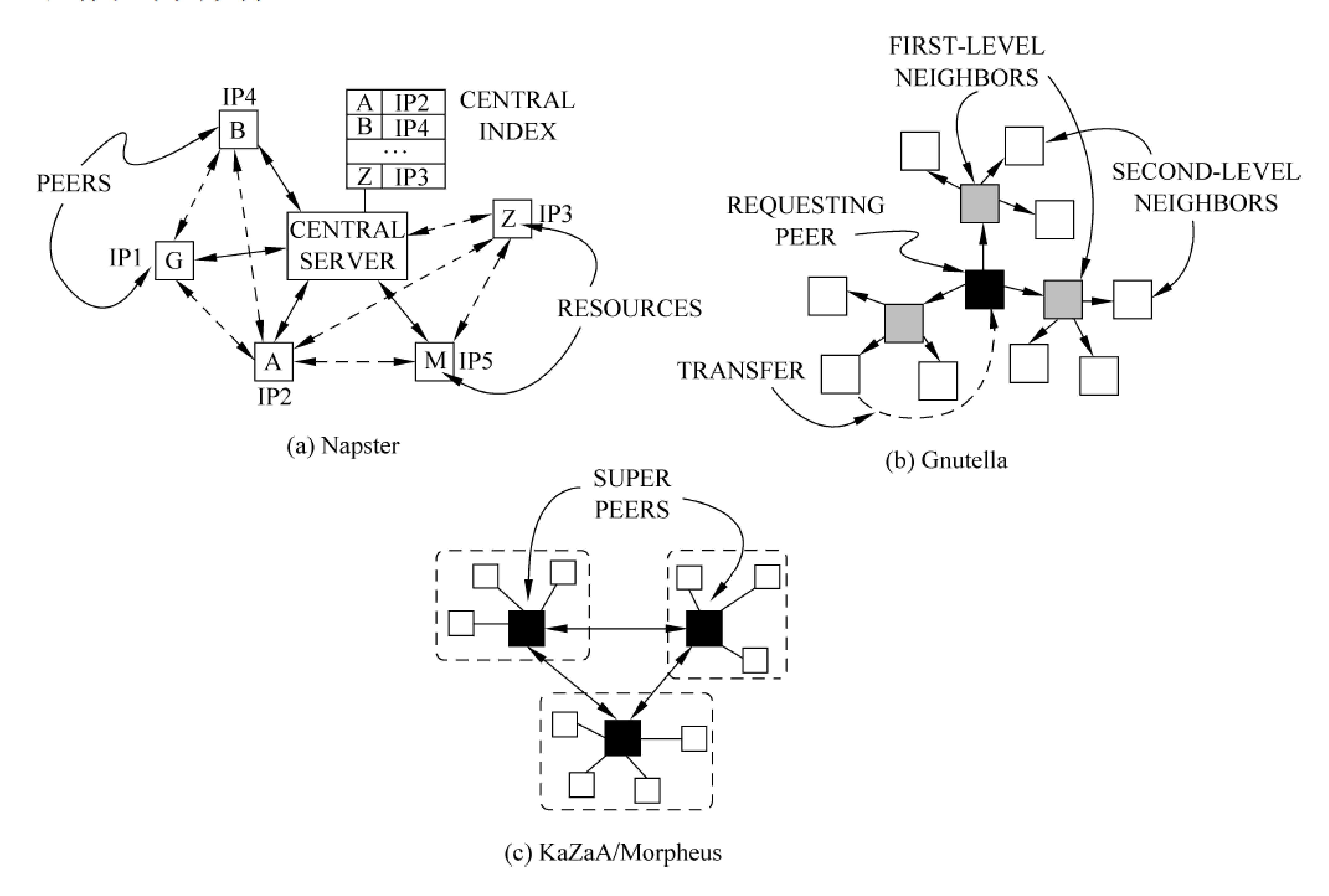


图 9-2　Napster、Gnutella 和 Morpheus/KaZaA 系统图

2. Gnutella

Gnutella[2]不仅文件传输是分布式的,资源发现也是分布式的。Gnutella 采用泛洪(flooding)定位资源,如图 9-2(b)所示。Gnutella 建立了一个网络,其中每个机器都连接到网络中的一部分其他机器,这些机器被称为邻居。例如,每台机器都知道它的邻居的 IP 地址并能够与它们进行通信。当机器发起查询时,它是通过向它的所有邻居集发送此查询请求,而它的邻居继续重复此操作向它们自己的邻居集发送查询请求,直至资源找到。这种设计避开了 Napster 的中心服务索引器,使得资源查找是分布式。然而,从技术角度来看,这并不是一种可扩展的方案。因为它对于每个请求而言都生成了太多的通信流量(communication

traffic)[114]。因此,泛洪需要采用 TTL(Time To Live)来控制流量。但是,这又产生一个现象是,尽管在 TTL 为 0 查询中止时的下一跳可能就是想要的资源,资源发现仍将失败。也就是说,对于资源确实是存在网络上的情况,由于资源处于请求节点的查询范围外(也许仅仅是在 TTL 查询中止的下一跳),它仍然是不可存取的。

3. Morpheus/KaZaA

Gnutella 的泛洪流量太大,系统如同 Morpheus[5] 和 KaZaA[3] 都通过采用超级节点[115] 来减轻泛洪流量,如图 9-2(c)所示。在这些系统中,本地节点形成一个组,并选出一个超级节点代表它们参与网络资源的共享。所有超级节点组成一个叠加网络。组内查询只需查询本组超级节点,仅当超级节点没有发现时,超级节点才将此请求传递给别的超级节点来进一步查询,由于查询具有本地性(locality),从而使得查询更加有效。超级节点是那些具有更高带宽、磁盘空间和处理能力的节点,通过它来缓存 meta-data 从而提高搜索效率。普通节点把其要共享的文件的 meta-data 上载到超级节点。在超级节点之间使用类似于 Gnutella 的广播搜索机制,由于超级节点的数量较少,所以网络流量不至于过大产生类似于 Gnutella 的不可扩展问题。但是这个方法仍然消耗大量带宽以使得超级节点维持本区域内的普通节点的 meta-data。因此尽管它修正了泛洪的一些问题,这种途径仍然是内在的不可扩展。

4. Freenet

Freenet[18] 是一个基于 Java 的跨平台分布式文件存储系统,相比于非结构化的其他系统,其最大的特点就是匿名。文件的发布者、查询者包括文件的持有者在 Freenet 中都是匿名的。为了实现匿名,Freenet 在路由上降低了效率,路由中的每个节点不能判断前一个节点是否是文件的请求者、也不能判断后一个节点是否是文件的持有者。

为了查询一个文件,用户必须首先得到或者计算对应此文件的二进制文件键(binary file key,通过 160 位的 SHA-1 哈希对文件描述串计算得到),然后发送包括文件键和 TTL 值的请求。当节点收到请求后,首先本地匹配,匹配成功则返回应答。如匹配失败,则在其路由表中查找与请求键最接近的键,并把请求递交到相应的节点。如果请求最终匹配成功且返回数据,数据将在请求路径上逆向转发,沿途每个节点在本地缓存该数据,并在路由表里创建关联实际数据源和该文件键的表项,以后对该文件的请求将从本地缓存中立刻得到满足。

如果由于目标节点失败或者查询路径循环出现导致某节点转发给下游节点的请求失败,该节点使用第二最近的文件键,然后第三最近文件键,依此类推。如果节点尝试了路由表中的所有节点均失败,就给上游节点报告转发失败消息,上游节点同样对路由表中各个节点进行尝试。如果到达 TTL 极限,文件请求失败消息向后返回到文件请求者并不再尝试。

Freenet 中查询文件的一个示意如图 9-3 所示。

Freenet 中没有明确实现某个具体节点负责某部分文档,查找采用的是搜索文件副本的方式,因此可以提供某种程度的匿名,但是不保证一定能找到在网络中存在的文档,并且降低了路由效率。

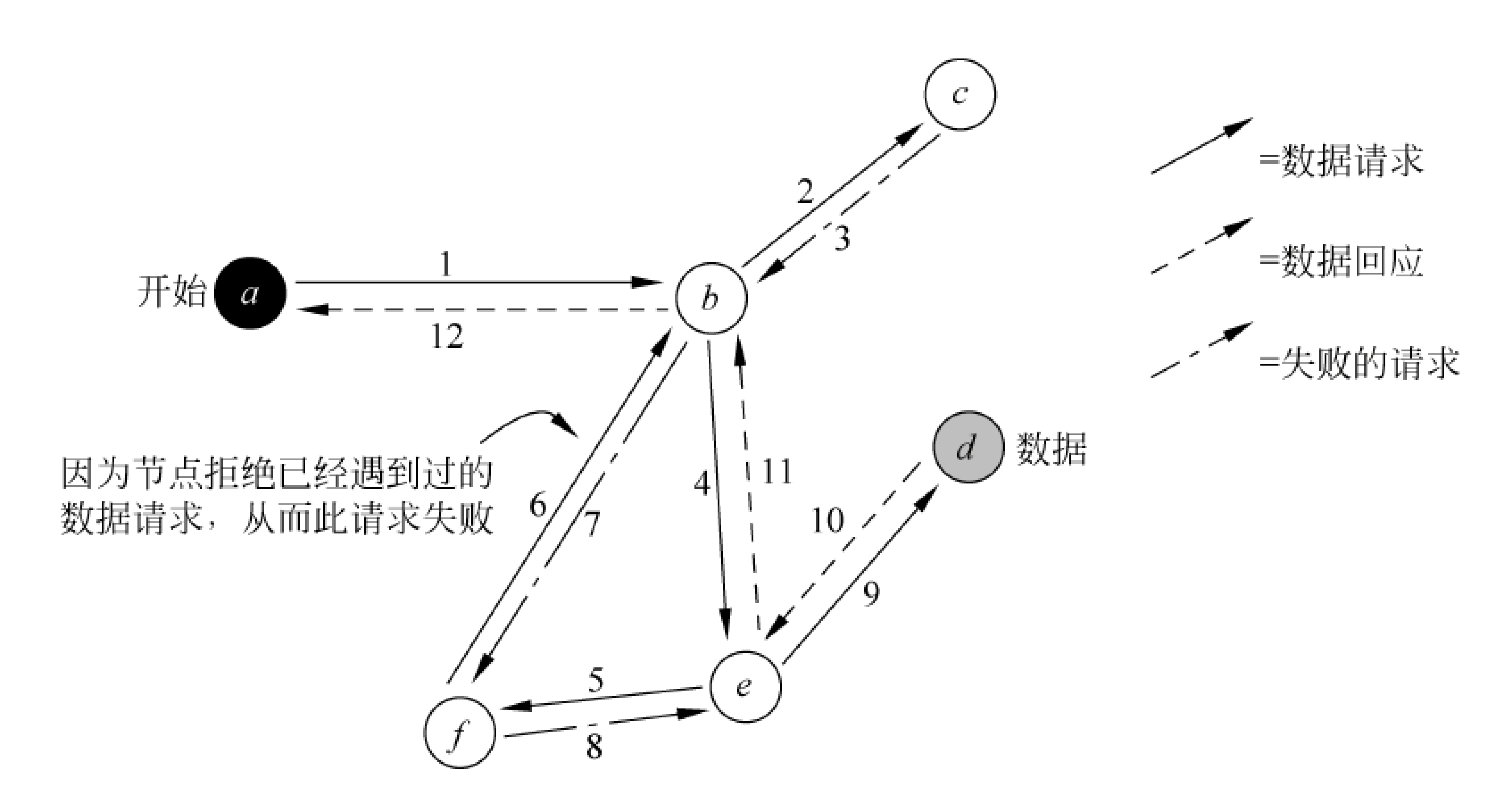


图 9-3　Freenet 中查询一个文件[18]

9.3.2　第二代 P2P 系统

本节介绍经典的 DHT-P2P 系统，突出它们的主要设计部分和功能特点。

1. CAN

伯克立和 AT&T 设计了基于 DHT 的查找和路由算法 CAN[22]。CAN 通过在现有的网络之上抽象出一叠加网(overlay network)，将其中所有节点映射到一个 n 维的笛卡儿空间中，并为每个节点尽可能均匀的分配一块区域(zone)，如图 9-4 所示。CAN 采用的哈希函数通过对(key，value)对中的 key 进行哈希运算，得到笛卡儿空间中的一个点，并将(key，value)对存储在拥有该点所在区域的节点内。CAN 采用的路由算法相当直接和简单，知道目标点的坐标后，就将请求传给当前节点邻居中坐标最接近目标点的节点。CAN 是一个具有良好可扩展性的系统，给定 n 个节点，系统维数为 d，则路由路径长度为 $O(n^{1/d})$，每节点状态信息和网络规模无关为 $O(d)$。

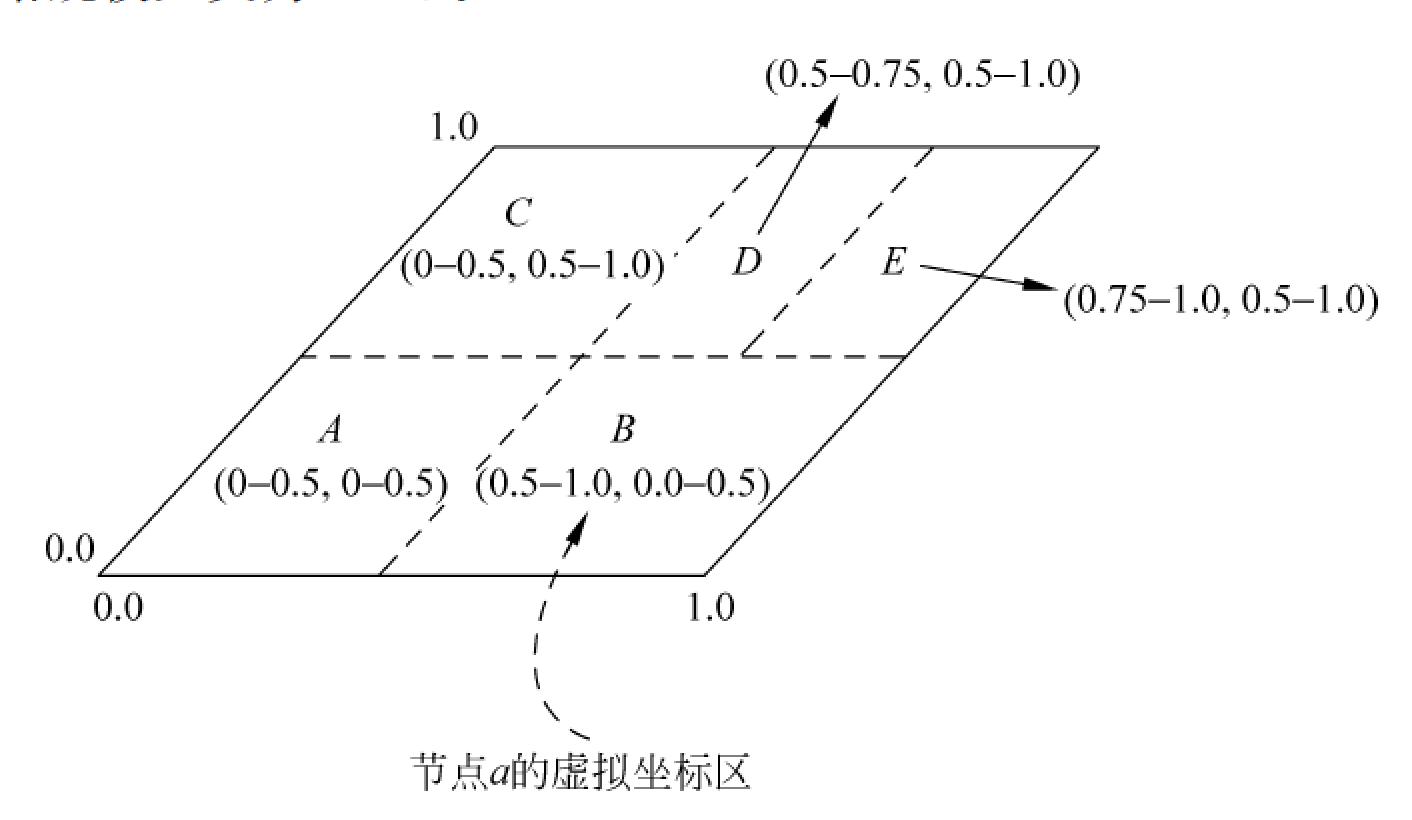


图 9-4　CAN 中拥有 5 个节点的二维空间拓扑[22]

一个新节点加入网络时，首先哈希该节点到叠加网坐标空间的一个点，然后拥有此坐标的区域分割为两半，并将其中一半分配给此新节点。分割的维向是在所有可能的维向中采

用一种固定的 round-robin 方式而选择的。失败节点的处理也采用同样的方式，也就是合并失败邻居的区域。不同的附加设计改善有：多坐标空间改善数据可用性，每个区域多个节点和多哈希函数改善容错等。

2. Tapestry

Tapestry[24] 源于 PRR[34] 路由机制。PRR 是一种支持叠加网络对象定位和消息路由的算法，但它基于理想的静态环境，并且没有提供很好的负载均衡。在 PRR 算法中，节点通过一次纠正一位的方式递交查找请求，所以节点匹配自己标识符的每一个前缀而下一位不同的邻居信息。对于 n 个节点的系统，每个节点有 $O(\log n)$ 个邻居，由于每跳纠正一位，所有路由路径为 $O(\log n)$ 跳。PRR 算法的路由过程如图 9-5 所示。

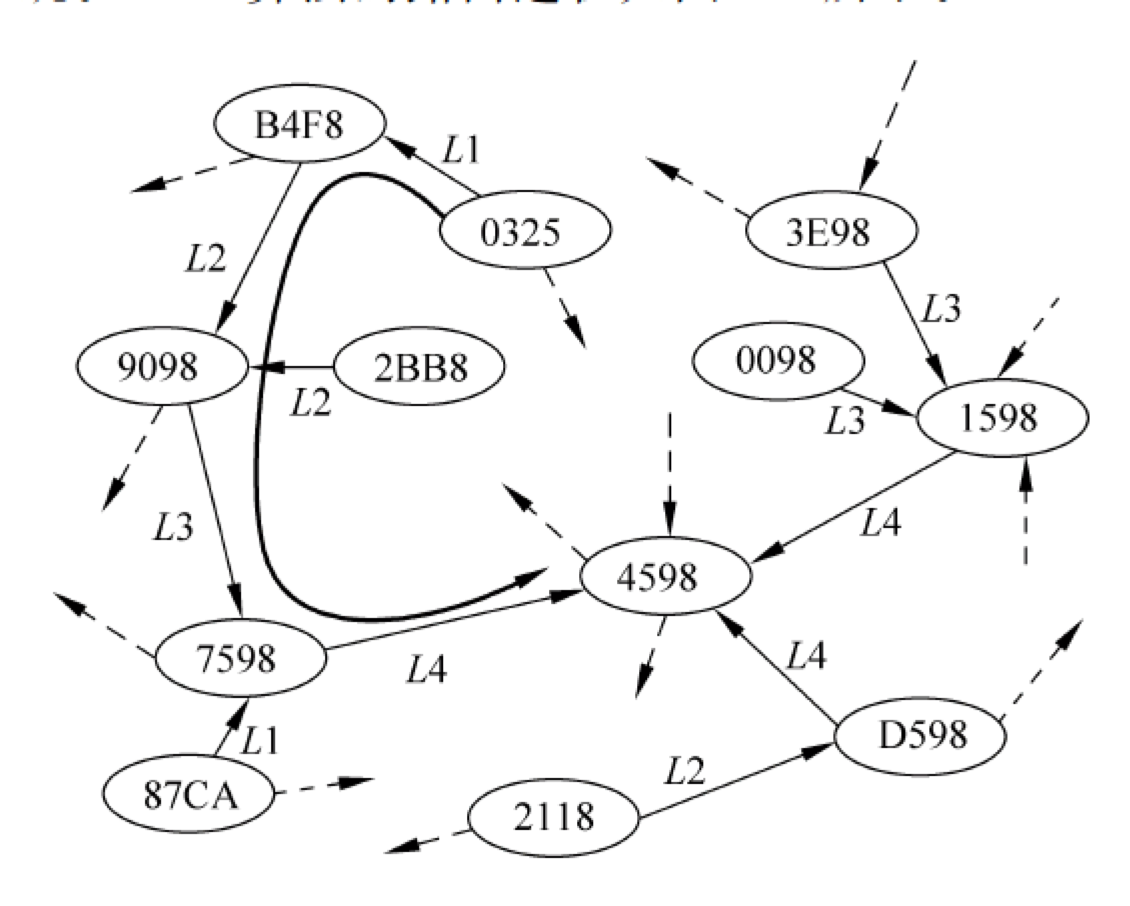


图 9-5　PRR 路由示例

注：节点 0325 到节点 4598 的路由过程由粗线标识，Tapesry 和 Pastry 都是基于 PRR 算法[34]

图 9-5 所示的路由算法具有如下特性：如果知道 n^2 个节点的延迟（或者是某种度量），节点能够选择邻居节点以最小化平均路径延迟且保证两节点间叠加路径延迟在物理路径延迟的常量因子之内。

PRR 服务器 S 通过路由消息给对象 O 的“根节点”声明它有对象 O。根节点是网络中的用来放置嵌入式树根的唯一节点。发布节点的过程包括发送消息给根节点，在路径上的每一跳，节点存放消息中的<Object-ID(O)，Server-ID(S)>映射信息。该映射信息是对服务器 S 上存放的对象 O 的指针，而不是对象 O 本身。当多个对象存在时，中间路径上节点只存放最近对象的映射。

在 PRR 的查询请求中，客户把对对象 O 的查询请求发给 O 的根节点，在每一跳，如果消息遇到包含对象 O 的映射信息的节点，立刻重定向到包含 O 的服务器。否则消息一直递交到根以确保找到对象 O 的映射。

Tapestry 为适应 P2P 网络的动态特性，做了很多改进。Tapestry 的路由和定位机制和 Plaxton 很相似，不同的是 Plaxton 中当有多个对象备份时节点将查询路由到距离其最近的根节点，而 Tapestry 的根节点存储了所有的备份的映射信息以提高语义灵活性。除此之外，为了适应动态的环境，Tapestry 增加了额外的机制实现了网络的软状态，并提供了自组织、鲁棒性、可扩展性和动态适应性，当网络高负载且有失效节点时候性能有限降低，消除了

对全局信息的依赖、根节点易失效性和弹性差的问题。

3. Pastry

Pastry[23]也是基于 PRR[34]路由机制，但不同于 Tapestry 的是它采用了基于后缀的路由代替了基于前缀路由。在 Pastry 中，每个节点分配一个 128 位的节点标识符号(node_id)，所有的节点标识符形成了一个环形的 node_id 空间，范围为 $0\sim2^{128}-1$，节点加入系统时通过哈希节点 IP 地址在 128 位 node_id 空间中随机分配。

Pastry 中节点的路由状态包括三部分：路由表 R、邻居集 M 和叶子集 L。

节点的路由表由$\lceil \log_{2^b} n \rceil$行，每行有 2^b-1 个入口的表项组成。行 n 的 2^b-1 个入口指向其 node_id 和当前节点的 node_id 共享前 n 位但第 $n+1$ 位不同的 2^b-1 个表项。值 b 的选择考虑了路由表的长度和路由跳数的权衡，b 越大则路由跳数少但需维护的路由信息多。

邻居集 M 包含了同本地节点最接近(根据 proximity metric)的$|M|$个节点，不用于路由转发而用于保证 locality 特性。

叶子集 L 是$\lceil L/2 \rceil$个其 node_id 最接近且大于本地节点 node_id 的节点和$\lceil L/2 \rceil$个其 node_id 最接近且小于本地节点 node_id 的节点集合，典型的$|L|$和$|M|$值是 2^b 和 2×2^b。

给定一 key，节点首先检查该 key 是否落在叶子集范围内，如是则直接把查询请求递交到叶子集中 node_id 最接近该关键字的节点，查询结束；如在叶子集范围之外，则节点把查询请求首先转发到其 node_id 和 key 共享的位数比本地 node_id 和 key 共享的位数至少长一位的节点，如果不存在该节点则转发到共享位数一样长但 node_id 比本地 node_id 数值上更接近 key 的节点。

给定 n 个节点，Pastry 叠加网络中，路由一个消息需要 $O(\log n)$步，每节点需要维持 $O(\log n)$个入口，而且 Pastry 路由具有 locality 特性。

4. Chord

麻省理工学院设计了一种分布式的可扩展的查找和路由协议 Chord[21]，Chord 通过将(key，value)对中的 key(如文件名)和网络节点分别进行哈希，并将哈希值映射在相同的值空间，将(key，value)对存储在最接近 key 的哈希值的节点上。一般来说，Chord 将 n 个节点哈希(采用一致性哈希[28])到具有 $\log n$ 位的标识环上。每个节点 x 存储指向它的直接后续 **Successor**(沿环顺时针方向的最近节点)。它也维护一个有 $\log n$ 个表项的指针表(**finger table**)。第 i 表项存储 $x+2^{i-1}$的后续标识。图 9-6 是一个有三个节点的 Chord 标识环示意图。Chord 路由算法采用了类似二分查找的方法，每次查找发送的消息数为 $O(\log n)$。稳定状态下，在一个 n 节点的系统中，每个节点需要维持 $O(\log n)$个其他节点信息，解析查找需要 $O(\log n)$个消息。当节点加入和离开系统时，Chord 为维持路由信息一致性最多发送 $O(\log_2 n)$个消息。当路由信息不一致时，性能只有限降低。在节点加入离开频繁的 P2P 网络，每个节点只需要更新少量信息就可保证正确的路由查询。为了容错，每个节点可维护一个后续链代替单个后续。动态节点加入与离开的性能分析可见[21]。Chord 的特点是简单、可以证明的协议正确性和良好的性能。

9.3.3 新型 DHT-P2P 系统简介

本节介绍基于 DHT 设计的新型 P2P 系统。这些设计大致展示了当前的最新技术和方

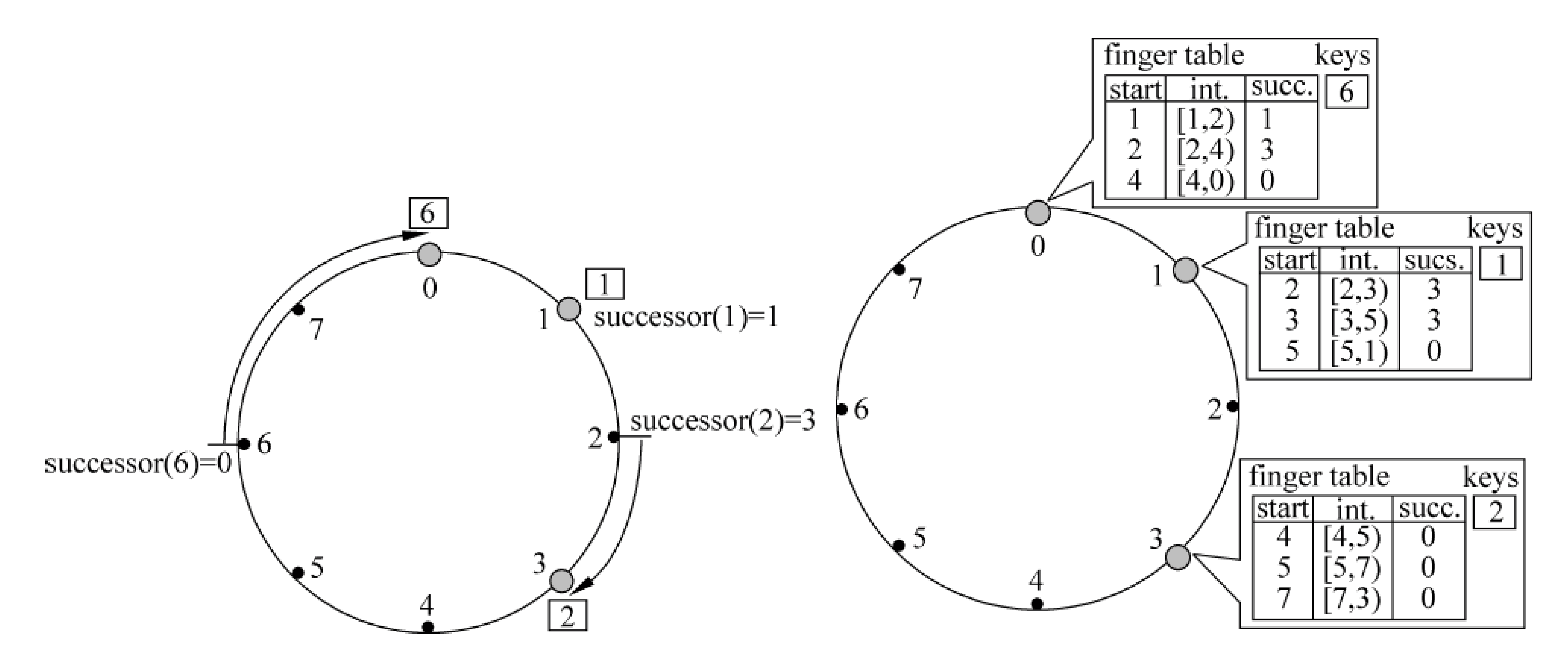


图 9-6　拥有三个节点 0、1 和 3 的 Chord 标识环

向。从拓扑结构特征可归为两类：

(1) 新型确定性拓扑构造；

(2) 随机化拓扑构造。

确定性与随机化的拓扑的主要区分指路由表中是否存在随机链构造。

1. 新型确定性拓扑构造

典型系统如 CAN、Chord、Pastry、Tapestry 都是确定性拓扑。在此之后，研究探索了将 DHT 布置于不同的拓扑图结构中，从而实现了“常量度路由效率”①及增强的一些特性。

1) Kademlia

Kademlia[29]设计与 Chord 极相似，仅有一点不同。它使用叠加网中两节点标识值的异或作为距离度量(metric)，这点使得任意两节点具有对称的距离，区分开 Chord 的不对称距离。另外，这种网络结构也使得每个节点能够从一定范围内的节点里选择邻居，从而网络更具有容错性，并且可选节点的增多也使系统能够更好地适应网络的动态改变。

2) P-Grid

P-Grid[33]分布数据在一个虚拟二叉树图结构。P-Grid 树中节点位置由节点在树中的路径决定。树中每个节点含有一个路由表，其中每一路由项指向同层中一个与该节点在其某路径位上与该节点具有相反值的节点。查询采用基于前缀的路由。P-Grid 提供一定的本地性(locality)支持，但由于层次性，节点存在负载不均衡。

3) SkipNet

SkipNet[116]利用了跳表(SkipList)图结构。所有节点以它所在的定位区域(domain)名被排序，然后以跳表方式形成统一的结构化拓扑。此外，在节点所属定位区域内的所有节点以 DHT 方式连接且资源一致性分布在此区域内。SkipNet 提供了两阶段搜索，首先是在名字空间以跳表方式搜索，然后在区域内以 DHT 方式搜索。SkipNet 有三种特性：路径本地化，内容本地化，有限负载均衡性。

① 常量度路由效率是指具有 $O(1)$ 常量度路由状态，但达到 $O(\log n)$ 网络路径长的路由效率。

4）Koorde

Koorde[32]将数据分布于 de Bruijn 图。每个节点有一个 t 位标识 $b_t b_{t-1} \cdots b_1$ 并连接到标识为 $b_{t-1} b_{t-2} \cdots b_1 0$ 和标识为 $b_{t-1} b_{t-2} \cdots b_1 1$ 的两节点。它提供了"常量度路由"，t 增大有利于改善容错。

2. 随机化拓扑构造

Viceroy 是第一个随机化 DHT 拓扑。随机化拓扑的特点是具有较好的拓扑动态适应性及更好的路由状态效率。随机化拓扑有三个不确定性源[35]：

① 节点总数不能够精确知道。

② 节点标识不是一致性均匀分布。

③ 不同节点有不同的随机选择。

1）Viceroy

Viceroy[31]将数据分布于一个 butterfly 网络。它的拓扑是一个具有前趋和后续两条基本链的双向环并且具有 5 条随机链（通过一定概率选择长链）。它具有常量度路由效率，即 $O(1)$链可达到 $O(\log n)$路径长。

2）Symphony

Symphony[30]简化了 Chord 环的链接（link），仅仅用两个基本链和一条随机链共三条链就代替了 Chord 环的 log n 条链。两个节点 u 和 v 间存在一条链接的概率与此两节点间的距离成反比。它实际形成一个 small world 网络，具有较优的路由效率，在 $2k+2$ 条链下，可达到 $O((\log^2 n)/k)$网络路径长。

9.3.4　比较与分析

本节对于前述 P2P 系统进行比较和分析。首先对于两代 P2P 系统归纳比较如表 9-2 所示。

表 9-2　两代对等系统性能比较

性能特征	第一代	第二代
路由模式	TTL 限制的泛洪	DHT
可扩展性	差	好
确定性搜索	不可	可
查询路数	不定	$O(\log n)$
空间开销	常量(3～7)	常量～$O(\log n)$

注：n 是系统内节点数量，log n 即 $\log_2 n$。

由表 9-2 可知，第一代系统采用泛洪路由机制，从而可扩展性差，查询具有不确定性；相比之下，第二代系统基于分布哈希表路由机制，具有可扩展性以及确定性的查询性能。此外，从表中也可以看到，第二代系统的空间开销相对而言比较大，因而常量度空间开销是所想要的。

接下来，对前述的 DHT-P2P 系统的路由性能进行比较（见表 9-3）。

表 9-3　不同 DHT-P2P 系统的性能比较

P2P 系统	拓扑结构	路由状态	路由路数
CAN	Torus	$O(d)$	$O(n^{1/d})$
Tapestry	Plaxton Tree	$O(\log n)$	$O(\log n)$
Pastry	Plaxton Tree	$O(\log n)$	$O(\log n)$
Chord	Ring	$O(\log n)$	$O(\log n)$
Kademlia	XOR Tree	$O(\log n)$	$O(\log n)$
Symphony	Randomlized Link	$2k+2$	$O((\log^2 n)/k)$
Viceroy	Butterfly	7	$O(\log n)$
Koorde	de Bruijn	3	$O(\log n)$
P-Grid	Virtual Binary Tree	$O(\log n)$	$O(\log n)$
SkipNet	Skip List	$2\log n$	$O(\log n)$

注：n 是系统内节点数，d 是 CAN 的维数，k 是 Symphony 的链接数，$\log n$ 即 $\log_2 n$。

如表 9-3 所示，DHT-P2P 系统具有“常量度路由效率”，即系统路由具有常量度路由状态及 $O(\log n)$的路由跳数的特性。我们可以从表中了解到，对于 DHT-P2P 系统设计而言，“常量度路由效率”取决于其拓扑设计。

(1) 采用特殊图结构如 Koorde。

(2) 采用随机化拓扑如 Viceroy。因而，研究可从这两方面进行更多的努力。此外，综合现有的 DHT-P2P 系统的性能分析，目前最好的路由状态效率(可参考 9.5.1 节)是在 $O(1)$状态下达到 $O(\log n)$效率，因此我们提出相应的一个开放问题：在 $O(1)$的低限状态上，超越 $O(\log n)$效率是否可行？各相关方面的性能又将如何？

9.4　DHT-P2P 的典型应用

P2P 系统有广泛的应用如普及计算[36]、协同工作[37]、即时通信[38]、资源共享[4]、分布存储[39]、信息检索[40]等。本节所想强调的是 DHT 技术的特别适用应用，所以仅选择出作者认为重要的而且具有 DHT 技术应用特色的领域。

9.4.1　广域网络存储

传统网络存储采用集群方式需要高容量高性能服务器，是一笔很大投资，而且缺乏本地特性，可扩展性受很大限制。P2P 技术提供了经济的、具有本地特性、灵活方便的可扩展功能，因而受到极大关注。特别是为了保障网络中存储的资源能够确定性查找，基于 DHT 的 P2P 系统方案是一个重要的选择。目前大部分 P2P 网络存储是基于 DHT 技术。下面介绍典型的网络存储应用。

OceanStore[39]提供了全球范围内的一致性数据存储，采用 Tapestry 作为底层的路由机制。它可以自动从服务器和网络的失效中恢复，并可以混合使用新的资源和适应不同的使用范式。PAST[7]是 Rice 大学和微软研究院的联合研究项目，它是大规模协同的分布文件存储，提供可扩展性、可用性、安全性和协同资源共享，采用 Pastry 作为它的路由机制。它的存储特性是文件所有者可以发布也可以收回文件的存储，提供的是一致性存储。CFS[41]

是 MIT 的研究项目,它提供了一致性的分布协同文件存储,采用 Chord 作为它的路由机制。从文件系统的可读写角度看,PAST 系统仅支持只读,CFS 基本上是只读文件系统,但也存在粗粒度的修改,即文件所有者可以修改,而其余只能读。因此,可读写的文件存储系统是进一步的研究。IVY[42]是 MIT 最近的项目成果,它是一个多用户可读写 P2P 文件系统,基于 Chord 通信子层,适合广域的小型协作组。IVY 使得这些组避开了集中式文件服务器所固有的可靠性和信任问题。一个 IVY 文件系统只由一组 log 组成,每一个用户对应一个 log。每一参与者通过协商所有 log 查询数据,但是修改则只是修改自己所有的 log。这种安排使得 IVY 可以不通过锁机制而维护元数据的一致性。当底层网络是满连接时,IVY 提供 NFS-like 传统语义,如 close-to-open 一致性;当网络分区时,DHASH 层的复制使得修改可在分区内进行,并借助每一 log 所具有的 version 实现分区合并时的一致性。Mnemosyne[43]是基于 Tapestry 的安全存储系统,存储的文件仅合法的用户能够存取它们,而攻击者不能够得到它们的内容。文件被加密存取且通过编码技术如 Erasure Code[44]保证了容错性。

9.4.2 网页发布和缓存

传统网页发布和缓存是由高性能服务器提供的,存在 flash crowd 问题[45]。P2P 技术可以通过缓存在各 peer 的缓存而减轻请求热点,并且使得网页可以在全球节点共同提供,减轻了对服务器的依赖。DHT 技术特性使得定位自主,不必依赖协调器的定位服务(如 YouServ[46]),有更大的可扩展性和灵活性。

Squirrel[47]是一个基于 Pastry 的分散式的 P2P 网页缓存。系统聚集所有本地缓存(cache)形成一个全球可扩展的网页缓存,而且并不需要附加的维护开销或者硬件投资。实验结果表明开发这样的一个系统不仅仅是可行的,而且比特定的网页服务器有更多的益处。

Coral[48]是纽约大学基于 DHT 的网页内容分发系统。节点有责任存储和复制网页内容。所有节点的缓存协同工作,任何时候都尽可能将所需要内容从邻近节点得到,从而使得提供网页的最初服务器上负载最小化以及尽可能减少了客户端的延迟。

9.4.3 组通信

由于 DHT-P2P 系统中采用连通图拓扑来组织节点,因而适合将它这种结构化特性用于组通信。

Pastry 已经被使用作为 Scribe[10]的基础。Scribe 是一个事件通知体系用于基于主题的出版/订阅系统。订阅者注册他感兴趣的主题并接收相关与这主题的事件。一个多播树用于维护每个相关于一个会合点(rendezvous,如叠加网上一个最接近主题标识的节点)的主题。每个树是通过加入从订阅者到特定主题的路由所组成。

DHT-P2P 系统也在应用层广播[49]以及应用层多播[9,50]得到应用,特别是应用层多播。应用层多播,顾名思义就是在应用层实现多播功能而不需要网络层的支持。这样就可以避免出现由于网络层迟迟不能部署对多播的支持而使多播应用难以进行的情况。应用层多播需要在参加的应用节点之间实现一个可扩展的,支持容错能力的叠加网络,而大规模哈希表查找机制正好为应用层多播的实现提供了良好的基础平台。当前有两种途径用于多播:智

能泛洪(intelligent flooding)和多播树(multicast tree)。前者用于基于 CAN 的 M-CAN[51]；后者用于 Scribe[9] 和 SplitStream[50]，它们都是基于 Pastry。两种途径在相同的工作负载下的比较显示多播树比泛洪途径性能优越[52]。

9.4.4 名字服务

DHT 机制通过哈希得到的键值对的方式及处理使得它非常适用于名字服务系统，使之具有高效、准确、高可扩展等特性。

简单分布式安全体系(Simple Distributed Security Infrastructure，SDSI)是一个已提出的公钥体系，其中名字定义在本地的名空间(namespace)并且长文件名能够链接多个名空间来代理使用认证的信任。ConChord[53] 是一个基于 Chord 的分布式 SDSI 认证目录服务，用于多名空间的名字解释和成员检验，检查一个认证对于一个特定的钥(key)是否可用。DHash[54] 把 Chord 作为提供 DNS 服务的一个可选服务结构，DNS 中的(主机名，地址)对通过 DHT 技术能够以分布式方式存储，从而有利于减轻 DNS 树结构查找的负载不均衡、根节点瓶颈等问题。

9.4.5 信息检索

搜索引擎是目前人们在网络中信息检索的主要工具。目前的搜索引擎如：Google、Yahoo 等都是集中式的搜索引擎。这种信息搜索方式是一种被动的搜索方式。用户不能够主动将自己的信息发布到搜索引擎上，也不能够保持搜索引擎所采集数据的实时性。P2P 技术能够提供一种主动的、实时性、高扩展性的信息检索，开辟了信息检索新方向。DHT 技术相比较其他非结构化具有在确定性使得信息能够以一种确定的高效方式定位，从而引起特别的关注。

文献[55]较为全面地讨论了构造 P2P 搜索引擎所要面对困难，并做出了可行性分析。特别是对于基于 DHT 技术的全文本网页搜索，提出了合理化的技术建议和一些创新的优化技术。PeerSearch[56] 是第一个具有分散化、确定性和非泛洪的 P2P 信息搜索系统，可基于内容和语义进行搜索。PIER[57] 提供了在 Internet 范围的 P2P 信息检索，能够提供基于关系查询的语义支持。

9.5 DHT-P2P 系统路由研究进展

路由算法是 P2P 系统的核心，算法的优劣直接关系到 P2P 系统的核心性能如可扩展性、可靠性、可用性等。Sylvia Ratnasamy[58] 等人在总结现有的 DHT-P2P 路由算法的基础之上提出了结构化对等网络面临的“十五个问题”。这些问题体现在五个方面：状态效率折中、容错性、路由热点、物理网络匹配和异构性。事实上，它引导了对于 DHT 路由的研究方向，研究进展分述如下。

9.5.1 状态效率折中

Ratnasamy Sylvia 等[58] 在 2002 年的 IPTPS 会议上介绍到当前 DHT 路由有两种模

式：(a)有一个路由表大小为 $O(\log n)$ 和网络路径长为 $O(\log n)$，比如 Tapestry、Pastry、Chord；(b)有一个路由表大小为 $O(1)$ 和网络路径长为 $O(n^{1/d})$，如 CAN。这两种模式事实上反映了路由表空间与路由的路径长度的一种互为消长，Ratnasamy Sylvi 等称之为路由状态效率折中(routing state-efficiency tradeoff)，其中状态指路由表所需要维护的邻居状态数目，效率指路由路径长度所代表的路由效率。一个重要的开放性问题是——是否能够实现结合以上两种模式优点的状态效率存在：$O(1)$ 的邻居数和 $O(\log n)$ 的路径长？这也被称为"常量度"路由问题。该问题引起了研究领域极大的兴趣，吸引了科研工作者对此的深入研究。最近已有一批常量度路由算法陆续被提出，这些算法设计利用了不同拓扑几何的特性如 de Bruijn[32]、Comb[59] 支持常量度路由。但是，Jun Xu 等[60] 分析也指出，拥塞可能是这些路由算法的突出问题。此外，不同的研究方法也涌现。如传统构造默认路由表的分布一致性(uniform)，然而文献[61]采用不规则路由表，根据节点的能力来划分路由责任对于状态效率的研究是一种较为新颖的方法。利用节点的社会链关系[62,63] 及随机路由算法[44] 也是状态效率研究的新方法。由于状态效率对于 P2P 系统的性能具有本质性影响，时至今日，状态效率折中的研究仍然是方兴未艾。

9.5.2 容错性

由于节点在 P2P 系统内自由加入和离开，因而容错性显得特别重要。一方面，系统要在面对路由节点失效或者离开时，能够仍然保证路由可行性和正确性；另一方面，在大规模失败时，网络可能由于节点间完全失去连接，形成孤岛现象，称为分区(partion)，这种分区现象也要处理好。对于前者的评估，Sylvia Ratnasamy 提出 static resilience 概念。static resilience 是指当节点失败并没有时间让别的节点重建别的邻居作为补偿时，也就是说邻居节点知道一个节点失败了但并不与别的节点建立任何新的邻居关系时，路由的可行性及效率。这个问题在文献[64]中得到了较好的回答。文献[64]认为 static resilience 与路由 Geometry 密切相关和环结构对于 static resilience 有效，此外还区分路由表的邻居为规则和持续两类，其中持续邻居对于 static resilience 更有效。相似的结论出现在文献[65]，不过，此时是将规则和持续这两类对应为长链和短链，并通过随机过程理论分析和实验得出短链对于 static resilience 有更重要的影响。文献[66]建立了考虑有一半系统内节点失效时系统的重建过程的分析模型，文献[67]利用 SkipNet 的特性考虑分区问题，但目前如何有效处理大规模失败的问题仍然是一个 open question。

9.5.3 路由热点

路由热点是指当存有资源的节点受到太多请求时，此节点的出入路由流量太大，引起所谓"热点"问题，即路由拥塞，节点由于负载过重而反应不过来。路由热点的解决方案一方面可以通过有效的复制和缓存策略[45,68]，以分流对于热点资源的请求，另一方面的重点是采用负载均衡技术。在 DHT 中，负载均衡可采用虚拟服务器(virtual server)的方法，根据节点的能力分配负载，如 Chord[21]。在文献[69]中基于 Chord 进一步考虑了虚拟服务器根据动态负载的情况进行迁移(transfer)的技术。其技术要点是重载节点将一些虚拟服务器迁移到轻载节点，并考虑了总体性能的均衡，而文献[70,71]则继续对于减少迁移的开销做了

一些改进和相应的理论分析。文献[72]采用的 the power of two 技术,在哈希发布文档时就将文档索引项采用多个哈希函数进行多个节点存储,并在查询时随机选取一个存储节点进行定位路由,从而较好地均衡了负载。

9.5.4 物理网络匹配

物理网络适应指的是叠加网络应尽可能与物理网络相匹配以减少节点通信开销和路由延迟。目前的研究主要是从两个互为补充方向进展:

(1) 利用 Internet 的层次信息如自治系统 AS、IP 前缀进行拓扑适应构造,这是一种直接与物理网络匹配的方法。

eCAN[73,74]借助 BGP 表来引导路由,使得节点通信本地化,减少通信开销和路由延迟。而文献[75]利用 IP 前缀使得路由充分利用 Internet 层次信息,并模拟路由层的最短路径算法以匹配物理网络。文献[76]提出一种 Sloppy 哈希技术以支持邻近路由。文献[77]则是根据节点的 IP 时延将叠加拓扑组织成多个层次,路由根据时延由短到长的层次进行,以尽可能减少整体通信时延。这些方式与物理网络的匹配性都是需要在设计时就考虑好,主要是基于叠加拓扑构造的方法。

(2) 采用路由选择技术,这是一种间接与物理网络匹配的方法。

DHT 路由选择技术有三种[78,79]。

① 邻近邻居选择(Proximity Neighbor Selection,PNS),即在路由表构造时,节点选择与该节点邻近(比如时延较短,物理位置靠近)的节点作为邻居。这种优化策略是能使路由跳转充分选择节点的物理邻近节点进行,从而使得路由间接与物理网络匹配。该技术依赖于节点在依此构造路由表时能够不影响总体路由跳数的自由度。在基于前缀的协议中(例如 Tapestry 和 Pastry),路由表的上层允许更自由的选择,而下层的选择自由度呈指数下跌。因此,第一跳的平均延迟非常低而随着每一跳指数增加,最后一条的延迟决定着整个查找过程的延迟。该类方法有好的延迟伸展、负载平衡和本地路由收敛特性[78],但是该类方法不支持如 CAN 和 Chord 这类要求在路由表中明确指出标识空间下一个节点标识的 DHT 算法。

② 邻近路由选择(Proximity Routing Selection,PRS),即在路由时,选择与该节点邻近的邻居作为下一跳。这种优化路由以匹配物理网络的构造技术是基于如果每步都是最短时延,那么总体时延也应该得到优化的原理。与 PNS 相反,它是在路由过程中进行的动态选择。假设每一跳都有 k 个节点可供选择,则每一跳的平均延迟可以从原来的网络中任意两个节点延迟的平均值减少到网络中任意一个节点到其他任意 k 个节点延迟最小值的平均值,所获得的性能提高同 k 的大小成正比,增大 k 的值意味着每个节点路由表的增大,从而需要更多的资源消耗以维持链接。除此之外,一味考虑选择延迟最低的节点转发查询也可能导致路由逻辑跳的数目变大。

③ 邻近标识选择(Proximity Identity Selection,PIR),即在新节点加入时,节点标识产生与节点所在的物理位置相关。它的原理是基于如果标识与物理布置相关,那么在以标识为基础的叠加网的路由就能够尽可能接近以 IP 为基础的物理网的路由。Landmark[22]技术广泛应用于 PIR。其方法是取 m 路标,然后每个加入的节点通过 ping 这 m 路标得到的值排列成 m 位,即为节点标识。由于此标识反映了与 m 路标的物理位置关系,因此,将标识与

物理位置结合起来。Brocade[81]是基于 PIR 的路由技术。eCAN[73]也采用了这种技术。文献[80]中实现 PIR 路由,它采用了一种基于 Landmark 的 binning scheme 技术来接近底层物理网的路由。然而,该方法有很多缺点:首先,它破坏了标识空间的均匀分布,在叠加网中造成了严重的负载不平衡问题。其次,由于映射算法的限制,该方法在一维空间中(Chord、Tapestry 和 Pastry)工作效果较差。再次,标识空间中相邻的节点由于物理也相邻,容易造成并发失败,而由于在 Chord 和 Pastry 这样的系统中节点在邻居中复制数据,因此也易造成安全和鲁棒性隐患。

以上三种方法中,PRS 是最轻量级的方法,但是性能受到 k 的限制,升高 k 值同时导致叠加拓扑的维护耗费更大。文献[64]实验试验结果表明,PNS 方法的性能优于 PRS,但受限于一些明确下一跳的 DHT 算法(如 Chord)。PIS 具有较好的平均每跳延迟,但是该种方法导致严重的负载不平衡并需要高维标识空间才能有效。

9.5.5 异构性

当前结构化 P2P 系统清晰或者不清晰地假定所有节点在资源(网络带宽、存储和 CPU)是一致分布的。消息在叠加网路由时并不考虑参与节点的能力差异。然而,测量研究表明 P2P 系统有极大的异构性[82],并且由于一部分节点非常有限的能力瓶颈可能引起路由算法失效。因此,将节点异构性考虑进去。利用异构性可以分配更多的能力给有高网络带宽、大存储容量和好的 CPU 处理能力的节点,这些节点也被称为超级节点。文献[83]中建立了一个超级节点虚拟层,并将本地节点组织成一个以超级节点为中心的组,并采用两阶段的路由过程,第一阶段是将请求路由给超级节点层,第二阶段时再将请求路由给超级节点所在组的目标节点。这种方式可以避开低能力节点的瓶颈,提高通信效率,降低通信延迟。文献[84]通过异构节点形成层次化 DHT 路由以得到与文献[83]相似的效果。

9.6 DHT-P2P 系统拓扑研究进展

拓扑对于 P2P 系统性能有重要的影响,直接相关于路由及查询性能。P2P 拓扑应该能够适应动态的网络环境的变化并与物理网络有更紧密的结合,增加 P2P 系统的适用范围和提高 P2P 系统的性能和效率。DHT-P2P 系统提供了结构化网络拓扑,目前的研究主要集中在以下三个方面展开:

(1) 控制拓扑维护开销:增强 DHT 拓扑的网络动态适应性。

(2) 层次化拓扑:克服 DHT 的平坦化结构,加入现实中的层次特征。

(3) 混合拓扑:结合结构化与非结构化拓扑的两种拓扑的优点。

这几方面的研究抓住了 DHT 技术本身的特点,将更好地提高 DHT 拓扑的适应性和效率。

9.6.1 控制拓扑维护开销

DHT-P2P 系统中由于将节点和文件紧密布置在一个结构化拓扑中,从而敏感于结构的动态变化。为了保障在动态网络环境下的正确结构和路由正常进行,维护拓扑的开销相

对比较大，特别是极动荡时可能超过系统的控制。

节点不断加入和离开的过程称为Churn，它对于系统性能有重要的影响。Sean Rhea等文献[85]研究了在Churn情况下性能的影响因子，并通过主动失败恢复、合理超时设置、邻近邻居选择等技术较好地控制了Churn。Bamboo[86]是基于这些技术的较好处理了Churn的一个DHT-P2P系统。文献[89]比较了Churn情况下不同的DHT(Tapesty、Chord、Kelips、Kademlia)的性能，得出在同样负载下，如果参数调整得足够好，它们都具有相似的性能。然而，参数的调整在不同的环境及协议下是极不同的，要具体情况具体分析。文献[87]建立了节点开销模型以用于观察系统的稳定性、热点及网络效率。文献[88]从路由表的构造上分析了开销减少的可能性。文献[90]分析了真实trace下的拓扑维护开销，并通过自反馈机制来调整性能，减少动态性对于系统性能的影响。文献[91]讨论了拓扑维护的现实性和意义。

9.6.2 层次化拓扑

DHT设计面向一个平坦(flat)的拓扑。它的优点是对于所有参与节点有全局一致的功能分布，并且没有单点故障。然而，层次性的缺乏使得特定系统的层次信息如目录层次结构、层次缓存等丢失，且管理较为困难。一些研究者提出了层次化拓扑的设计思想。Canon[92]是基于DHT的一种层次化拓扑设计，不仅保持原有DHT的状态路由效率，而且具有以下5点特性：

(1) 错误隔离；

(2) 用于多播的有效缓存和有效带宽；

(3) 匹配物理网络的拓扑；

(4) 层次化内容存储；

(5) 层次化存取控制。

TerraDir[93]提供了层次名空间(如DNS或UNIX文件系统名)的目录查询。它组织节点为层次树结构，父节点维护它的子节点的信息。为了容错和减少查询延迟，缓存和复制在系统中大量使用。文献[94]提出了一个通用的框架用于P2P叠加网络的层次组织。聚簇是实现层次化的常用方式。文献[75]是根据网络IP前缀进行分层次聚集的层次化DHT设计。文献[95]的观点是当前网络物理拓扑结构是层次化，因而设计层次化DHT以及更好的匹配物理网络和利用本地的局部特性。

9.6.3 混合拓扑

混合拓扑的研究是新兴的一个研究点。因为结构化和非结构化各有其优缺点，因而设计混合网络拓扑以综合两者性能优点，减轻其性能缺点，是一种有益的探索。Yapper[96]有一个与Gnuetlla相似的非结构的网络拓扑，但是引进哈希来存储数据键。每节点x被分配b种颜色：color(x)=hash(IP(x)) mod b。每个节点维护它的直接邻居(距离节点为一常量值的路由跳数c的节点)信息。资源存储在颜色与资源键哈希值(为一颜色)匹配的节点上，因而查找可限定在同种颜色的节点上，从而查询性能得到较大的改善。Kelips[99]在DHT基础上引入集中式组管理。其中n节点通过哈希聚簇O(sqrt(n))仿射组。每个组内有一超

级节点维护组内所有信息并由别的组进行通信。每个节点的空间占用是 $O(\mathrm{sqrt}(n))$，路由需要 $O(1)$时间。网络更新消息采用 gossip 协议[97]传播。不过，在超大规模的 P2P 系统里，这种组维护开销可能过大。LOO 等[98]在 IPTPS04 会议提出一种新的设计思想是：结构化适合定位查找稀罕数据项(rare item)，而非结构化适合定位查找流行数据项(popular item)，因而设计混合拓扑模型。网络中一部分为结构化 DHT 构造采用 DHT 方式定位索引稀罕数据项，另一部分为非结构化 Gnutella 构造采用泛洪方式定位通用数据项。LOO 等的实验表明此种混合方式使得查询效率和可扩展性都得到较好的提高。

9.7　DHT-P2P 系统查询研究进展

查询是 P2P 系统中的一个关键问题，也是 P2P 系统最广泛的应用。文献[100]对于 P2P 查询的问题做了一个总结，不过目标是在非结构化系统。对于结构化系统，它具有与非结构化很不相同的问题。最根本的原因是由于 DHT 采用哈希技术仅提供精确查询匹配，使得 DHT 查询受到极大的约束。DHT 查询的前景展望可参考文献[101]。

DHT 查询的研究主要从以下几个方面展开。它们的目的是增强 P2P 查询能力，从而扩展 P2P 系统的应用范围。

(1) 多关键字查询；

(2) 模糊关键字查询；

(3) 复杂查询。

9.7.1　多关键字查询

当前基于 DHT 的 P2P 系统为了能够支持多关键字的数据检索，一般首先对文档做索引，索引方式一般采用逆序索引(Inverted Index)，然后针对每个关键字将(关键字，文档标识)插入到 DHT 叠加网中。在 P2P 网络中的逆序索引分布如图 9-7 所示。

当进行多关键字组合查询时，针对每个关键字，根据 DHT 算法映射到存储该关键字的相应节点，在该节点上查询包含该关键字的文档列表，然后在节点之间顺序传送文档列表并计算交集。显然这种检索机制需要在网络上传输大量的中间文档列表数据，产生大量的网络带宽消耗，当前有许多研究集中在如何减少带宽消耗以及提高检索效率。在减少带宽消耗方面，文献[102]采用了 Bloom Filters[103]机制，如图 9-8 所示。例如，进行组合查询 $A\cap B$，设关键字 A 映射到节点 A，关键字 B 映射到节点 B。查询请求首先发往节点 A 检索包含关键字 A 的文档集合。在节点 A 上检索到包含关键字 A 的文档集合，然后计算该集合的 Bloom Filter 并将结果 $F(A)$发往节点 B，节点 B 检索本地得到包含关键字 B 的文档列表，并利用 $F(A)$计算出满足 Bloom Filter 测试的 B 文档集，记为 $F(A)\cap B$，该结果集再传递回给节点 A，并计算和 A 文档集的交集，将结果发送回客户端。

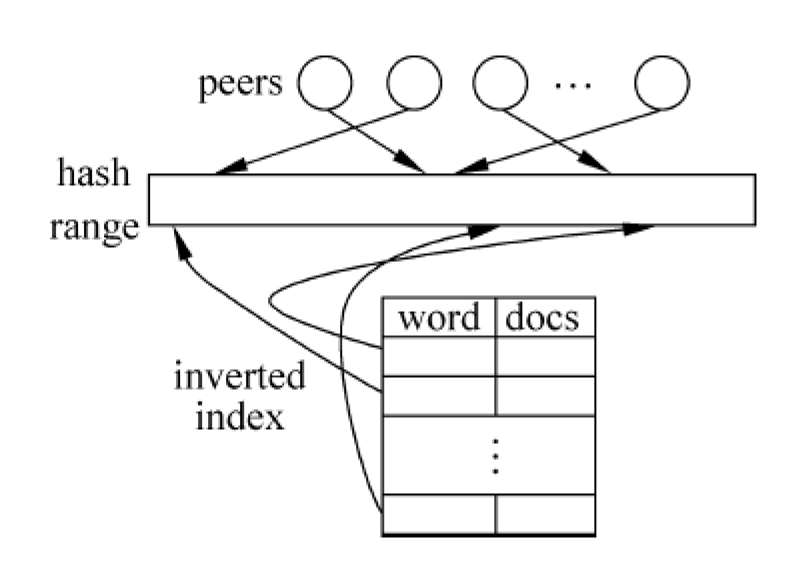


图 9-7　P2P 网络中的逆序索引分布[102]

如果用户数据检索的要求允许一定的误差，在上述过程中，节点 B 可以直接将 $F(A)\cap$

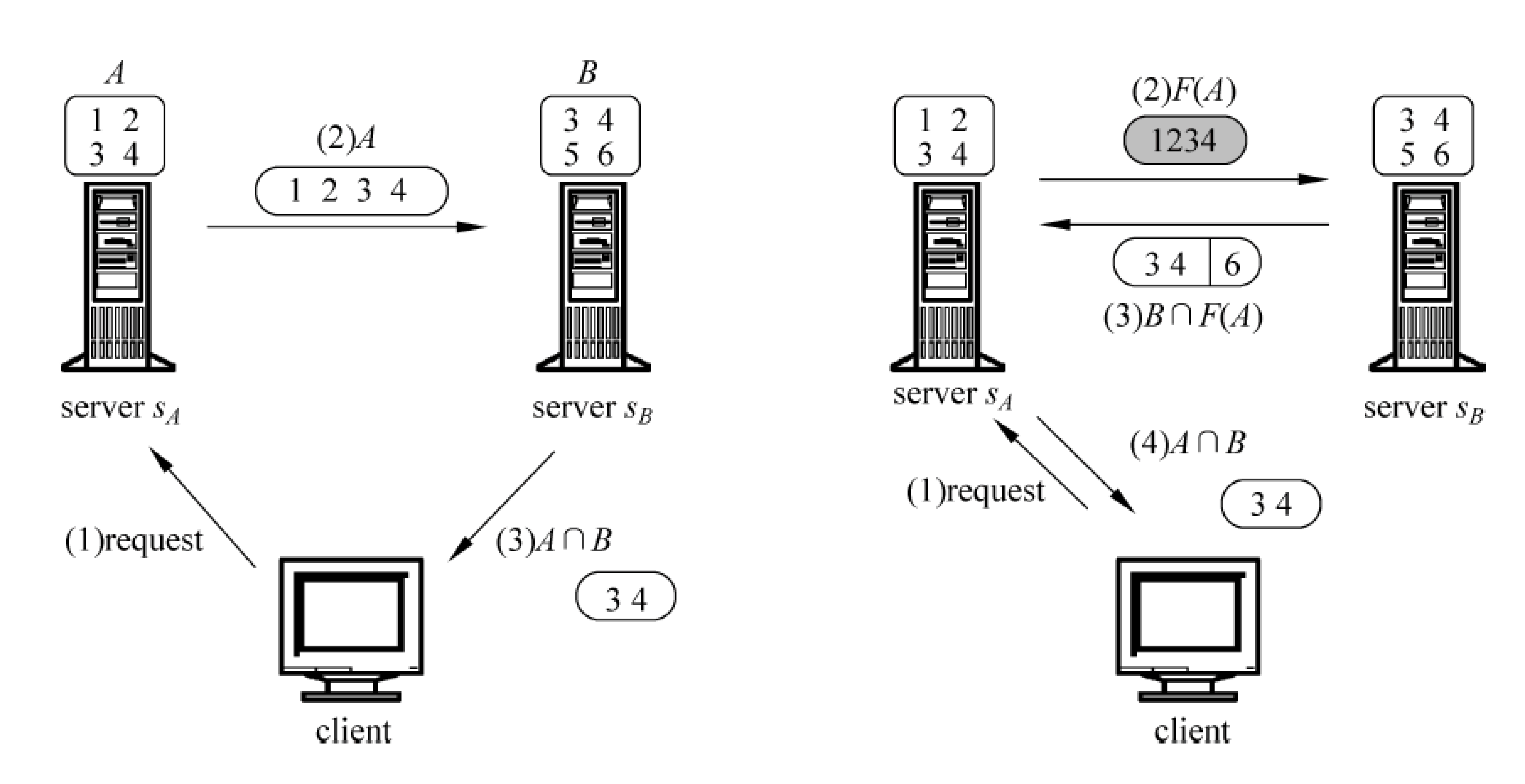


图 9-8 使用 Bloom Filter 计算组合查询 $A\cap B$[102]

B 发送给查询客户端,这可能会包含一些“正向错误”(false positive,即多包含了一些不正确的结果),Bloom Filter 的位数越多,正向错误率越低,但消耗的带宽越高。

多关键字查询中,一个关键点是减少传输占用带宽问题。如前面所介绍,Bloom Filter 是常用的技术。此外,还有一些可采用的技术。文献[104]利用了每次查询后的结果缓存(result-caching)技术来避免重复查询和减轻通信流量。比如,三个索引 a_i, a_j, a_k,查询 $a_i \wedge a_j \wedge a_k$ 直接从索引 a_i, a_j, a_k 得到显然比利用先前 $a_i \wedge a_j$ 的查询结果缓存再与 a_k 的交集计算得到所花费的代价大得多。文献[105]提出了 view tree 来存储和查询以前的结果缓存。此 view tree 实质是一个 trie 结构,据此能够很好地减少多关键字的查询开销。

以下这些技术并不直接相关于多关键字查询,但是其思想可以为多关键字查询所借鉴。

文献[106]提出部分查询技术以及相应的内容模式。其思想是考虑用户查询时并不需要所有副本,而是需要一些就够了,如查询一个流行歌曲只需要联系两三个网站就够了。因而,相对于传统查询返回所有结果集,部分查询能够有助开销减少和查询效率的提高。特别是它不敏感于流行 key,能够有效减轻热点问题。文献[107]提出轻索引技术,其思想是考虑建立索引的目的是为了有效进行查找,但并不是所有的内容都要索引,因而仅索引那些根据反馈计算得出值得索引的那些内容,从而降低系统的索引维护开销。部分查询技术和轻索引技术均能够有助于多关键字查询中的开销的减少,并有望提高查询效率。

9.7.2 模糊关键字查询

为使基于 DHT 的 P2P 系统能支持模糊查询,Harren 等[101]提出了一种采用 n-grams 的方法解决模糊查询问题,如 thoven 可分解为 tho、hov、ove 和 ven 等 3-grams。通过针对不同的值 n,可以分别建立索引。这种方法可以支持模糊查询,但是,它占用存储容量、网络带宽及处理开销都太大,扩展性很差。目前仍然没有很好的方案来解决这个问题,还有不少挑战性的工作需要去做。

9.7.3　复杂查询

应用查询应不只是仅仅支持关键字查询，还应该可以支持如同关系数据库的 SQL 查询（选择<Selection>、投影<Projection>、联合<Join>以及聚集<Group-by/Aggregate>）以及语义支持的 IR 模型查询等复杂查询。

文献[101]首次提出了 DHT 上 SQL 查询的设计思想，但仅实现了联合（join）。PIER[57]实现了一个 SQL 查询子集，但是在使用中经常会出现查询结果分布不均匀，出现大量的“热区”。作为 SQL 查询初步，范围查询是目前 SQL 查询的热点研究。文献[108]基于二维 CAN 提出了一种范围查找的方法，它利用了 CAN 的空间位置信息来存储查询结果集和并由此支持后续的子范围查询。Gupta 等[109]使用位置敏感哈希（locality-sensitive hashing）[110]来支持近似（approximate）范围查询。位置敏感哈希 LSH 是指哈希函数集合 H，$\forall h \in H$，有 $\Pr[h(A)=h(B)]=\mathrm{sim}(A,B)$。其中 A 和 B 为两集合，$\mathrm{sim}(A,B)$代表此两集合的相似度。它们用位置敏感哈希代替了 Chord 中的一致性哈希从而使得相似的数据项能够在哈希空间也接近。查询一些数据项将产生相似于被请求项的结果，因而支持范围查询。但是由于采用了不同的哈希功能代替一致性哈希，系统的负载均衡受到了影响。

文献[111]提出元数据搜索层（meta data search layer）用于统一数据和文件的描述和定义。事实上，这种建立元数据信息描述的手段已经在 Web Service 和服务发现机制中得到较为广泛的研究和应用。元数据搜索具有比简单数据搜索更丰富的语义。SOMO[112]借助于元数据来管理网络资源。Arturo Crespo 等[113]引入了路由索引（routing indices）的概念去允许邻居节点传递查询请求到更可能回答此请求的节点。如果一个节点不能够回答查询，它传递此查询到基于本地路由索引的一个邻居节点集，而不是随机选择邻居节点或者泛洪此请求给所有邻居节点。借助于路由索引，查询性能得到较好的提升，但它也需要额外的存储空间开销。

9.8　小结

本章首先对对等网络的定义及分类进行了介绍，然后针对对等网络的路由、拓扑和查询这三方面的研究工作进行了重点阐述，综述了当前对等网络的路由、拓扑和查询等方面的研究思路以及研究进展。

参考文献

[1]　Napster. http://www.napster.com.
[2]　Gnutella. http://gnutella.wego.com.
[3]　KaZaA. http://www.kazaa.com.
[4]　Limewire. http://www.limewire.com.
[5]　Morpheus. http://www.musiccity.com.
[6]　John Kubiatowicz, David Bindel, Yan Chen, et al. OceanStore: An Architecture for Global-Scale Persistent Storage. ACM SIGPLAN Notices, November 2000, 31(1): 190-201.
[7]　Rowstron A, Druschel P. Storage management and caching in PAST, a large-scale, persistent peer-to-

peer storage utility. In Proc. of SOSP (Oct 2001), ACM, 188-201.

[8] Bolosky W J, John R Douceur, David Ely, et al. Feasibility of a Serverless Distributed File System Deployed on an Existing Set of Desktop PCs. In Proceedings of the ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, Santa Clara, CA, USA, June 2000: 34-43.

[9] Miguel Castro, Peter Druschel, Anne-Marie Kermarrec, et al. SCRIBE: A large-scale and decentralized application-level multicast infrastructure. IEEE Journal on Selected Areas in Communications (JSAC) (Special issue on Network Support for Multicast Communications), October 2002, 20(8): 1489-1499.

[10] Rowstron A, Kermarrec A M. Druschel P, et al. SCRIBE: The design of a large-scale event notication infrastructure. In Proceedings of the Third International Workshop on Networked Group Communication (NGC), Lecture Notes in Computer Science, UCL, London: November 2001, 2233: 30-33.

[11] ThreeDegrees. http://www. threedegrees. com.

[12] Karl Aberer, Manfred Hauswirth. Peer-to-Peer Information Systems: Concepts and Models, State-of-the-Art, and Future Systems. Tutorial at The 18th International Conference on Data Engineering (ICDE), San Jose, California, 2002.

[13] Mike Miller. Discovering P2P. Sybex International, November 2001. ISBN-0782140181.

[14] Peer-to-Peer Working Group Committees. See http://peer-to-peerwg. org.

[15] Stephanos And routsellis-Theotokis, Diomidis Spinellis. A Survey of Peer-to-Peer File Sharing Technologies. White paper, Electornic Trading Research Unit (ELTRUN), Athens University for Economics and Business, 2002. http://www. eltrun. aueb. gr/whitepapers.

[16] Lv Q, Cao P, Cohen E, et al. Search and Replication in Unstructured Peer-to-Peer Networks. In Scott T. Leutenegger, editor, Proceedings of the 2002 International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS-02), SIGMETRICS Performance Evaluation Review, New York: 2002. 30(1): 258-259.

[17] Oram A. Peer-To-Peer: Harnessing the Power of Disruptive Technologies. O'Reilly, first edition, March 2001. ISBN: 0-596-00110-X.

[18] Clarke I, Sandberg O, Wiley B, et al. Freenet: A distributed anonymous information storage and retrieval system. In Proceedings of the ICSI Workshop on Design Issues in Anonymity and Unobservability (Berkeley, California, June 2000).

[19] Markatos E P. Tracing a Large-Scale Peer to Peer System: An Hour in the Life of Gnutella. In The Second International Symposium on Cluster Computing and the Grid, 2002. http://www. ccgrid. org/ccgrid2002.

[20] Petar Maymounkov, David Mazires. Kademlia: A Peer-to-peer Information System Based on the XOR Metric. In The 38 BIBLIOGRAPHY 1st International Workshop on Peer-to-Peer Systems (IPTPS'02), 2002.

[21] Stoica I, Morris R, Karger D, et al. Chord: A scalable peer-to-peer lookup service for internet applications. In Proc of SIGCOMM, ACM, Aug 2001: 149-160.

[22] Ratnasamy S, Francis P, Handley M, et al. A Scalable Content-Addressable Network. In Proc of SIGCOMM, ACM, Aug 2001: 161-172.

[23] Rowstron A, Druschel P. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. In Proc. of Middleware, ACM, Nov 2001: 329-350.

[24] Zhao B Y, Kubiatowicz J D, Joseph A D. Tapestry: An infrastructure for fault-tolerant wide-area location and routing. Tech. Rep. CSD-01-1141, U. C. Berkeley, Apr 2001.

[25] IRIS Project. http://project-iris. net.

[26] Litwin W, Neimat M A, Schneider D A. LH *—A scalable, distributed data structure. ACM Transactions on Database Systems, 1996, 21(4): 480-525.

[27] Gribble S D, Brewer E A, Hellerstein J M, et al. Scalable, distributed data structures for Internet service construction. Proc. 4th Symposium on Operating System Design and Implementation, OSDI 2000, 319-332.

[28] Karger David, Lehman Eric, Leighton Tom, et al. Consistent Hashing and Random Trees: Distributed Caching Protocols for Relieving Hot Spots on the World Wide Web. In Proceedings of the Twenty-Ninth ACM Symposium on Theory of Computing (STOC), El Paso, TX, May 1997: 654-663.

[29] Maymounkov P, Mazieres D. Kademlia: A peer-to-peer information system based on the XOR metric. In Proc. of IPTPS, Mar 2002: 53-65.

[30] Manku G S, Bawa M, Raghavan P. Symphony: Distributed Hashing in a Small World. In Proceedings of the Fourth USENIX Symposium on Internet Technologies and Systems (USITS), Seattle, WA, March 2003: 127-140.

[31] Dahlia Malkhi, Moni Naor, David Ratajczak. Viceroy: A Scalable and Dynamic Emulation of the Butter In Proceedings of the Twenty-First ACM Symposium on Principles of Distributed Computing (PODC), pages Monterey, CA, July 2002: 183-192.

[32] Kaashoek M F, Karger D R. Koorde: A Simple Degree-optimal Distributed Hash Table. In Proceedings of the Second International Workshop on Peer-to-Peer Systems (IPTPS), Berkeley, CA, February 2003.

[33] Karl Aberer, Philippe Cudré-Mauroux, Anwitaman Datta, Zoran Despotovic, Manfred Hauswirth, Magdalena Punceva, Roman Schmidt. P-Grid: A Self-organizing Structured P2P System. ACM SIGMOD Record, Sep 2003, 32(3).

[34] Plaxton C G, Rajamohan Rajaraman, Richa A W. Accessing Nearby Copies of Replicated Objects in a Distributed Environment. Theory of Computing Systems, 1999, 32(3): 241-280.

[35] Manku G S. Routing Networks for Distributed Hash Tables. In Proc. 22nd ACM Symposium on Principles of Distributed Computing, PODC 2003, June 2003, 133-142.

[36] SETI@Home. http://setiathome. ssl. berkeley. edu.

[37] Groove. http://www. groove. net.

[38] Jabber. http://jabber. org.

[39] Kubiatowicz J, Bindel D, Chen Y, et al. OceanStore: An Architecture for Global-Scale Persistent Storage. In Proc. of ASPLOS (Nov 2000), ACM.

[40] Steve Waterhouse. JXTA Search: Distributed Search for Distributed Networks, Whitepaper, Sun Microsystems, Palo Alto, Calif, 2001; In http://search. jxta. org/protocol. html.

[41] Dabek F, Kaashoek M F, Karger D, et al. Wide-area cooperative storage with CFS. In Proc. of SOSP, ACM, Oct 2001: 202-215.

[42] Muthitacharoen A, Morris R, Gil T M, et al. Ivy: A read/write Peer-to-Peer file System. In Proc of OSDI, ACM, Dec 2002: 31-44.

[43] Hand S, Roscoe T. Mnemosyne: Peer-to-Peer Steganographic Storage. In Proc. of IPTPS, Mar 2002: 130-140.

[44] Hakim Weatherspoon, John D Kubiatowicz, Erasure Coding vs. Replication: A Quantitative Comparison, Electronic Proceedings for the 1st International Workshop on Peer-to-Peer Systems, IPTPS 2002.

[45] Tyron Stading, Petros Maniatis, Mary Baker. Peer-to-Peer Caching Schemes to Address Flash Crowds Crowds, 1st International Peer To Peer Systems Workshop, IPTPS 2002.

[46] Roberto J Bayardo Jr，Rakesh Agrawal，et al. YouServ：A Web-Hosting and Content Sharing Tool for the Masses. In Proceedings of the Eleventh International World Wide Web Conference(WWW)，Honolulu，HI，May 2002：345-354.

[47] Sitaram Iyer，Antony Rowstron，Peter Druschel. SQUIRREL：A decentralized，peer-to-peer web cache. In Proceedings of the Twenty-First ACM Symposium on Principles of Distributed Computing (PODC)，Monterey，CA，July 2002：213-222.

[48] Freedman M J，Freudenthal E，Mazires D. Democratizing Content Publication With Coral. In Proc of NSDI，March 2004.

[49] Sameh El-Ansary，Luc Onana Alima，Per Brand，et al. Efficient Broadcast in Structured P2P Networks. In Proceedings of the Second International Workshop on Peer-to-Peer Systems (IPTPS)，Berkeley，CA，USA，February 2003.

[50] Miguel Castro，Peter Druschel，Anne-Marie Kermarrec，et al. SplitStream：High-Bandwidth Content Distribution in a Cooperative Environment. In Proceedings of the Second International Workshop on Peer-to-Peer Systems (IPTPS)，Berkeley，CA，USA，February 2003.

[51] Ratnasamy S，Handley M，Karp R，et al. Application level Multicast using Content-Addressable Networks. In Proceedings of the Third International Workshop on Networked Group Communication (NGC)，Lecture Notes in Computer Science，UCL，London：November 2001，2233：14-29.

[52] Miguel Castro，Michael B Jones，Anne-Marie Kermarrec，et al. An Evaluation of Scalable Application-level Multicast Built Using Peer-to-Peer Overlays. In Proceedings of the Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)，San Francisco，CA，USA，March 2003.

[53] Sameer Ajmani，Dwaine E Clarke，Chuang-Hue Moh，et al. ConChord：Cooperative SDSI Certicate Storage and Name Resolution. Lecture Notes In Computer Science，Cambridge，MA，USA，March 2002，2429：141-154.

[54] Cox R，Muthitacharoen A，Morris R T. Serving DNS using a Peer-to-Peer Lookup Service. Lecture Notes in Computer Science，Cambridge，MA，March 2002，2429：155-165.

[55] Li J，Loo B T，Hellerstein J，et al. On the Feasibility of Peer-to-Peer Web Indexing and Search. In Proceedings of the Second International Workshop on Peer-to-Peer Systems (IPTPS)，Berkeley，CA，USA，February 2003.

[56] Tang C，Xu Z，Mahalingam M. pSearch：Information Retrieval in Structured Overlays. In HotNets-I，October 2002. Expanded version available as HP technical report HPL-2002-198，Peer Search：Efficient Information Retrieval in Peer-to-Peer Networks.

[57] Huebsch R J，Hellerstein M，Lanham N，et al. Shenker S. and Stoica I. Querying the Internet with PIER. In Proc 19th VLDB，Sep 2003.

[58] Ratnasamy S，Shenker S，Stoica I. Routing Algorithms for DHTs：Some Open Questions. Electronic Proceedings for the 1st International Workshop on Peer-to-Peer Systems，IPTPS 2002.

[59] Considine J，Florio T. Scalable peer-to-peer indexing with constant state. Tech. rep，CS Department，Boston University，September 2002.

[60] Xu J，Kumar A，Yu X. On the Fundamental Tradeoffs between Routing Table Size and Network Diameter in Peer-to-Peer Networks，IEEE Journal on Selected Areas in Communications，Jan 2004，22(1)：151-163.

[61] Hu Jingfeng，Li Ming，Zheng Weimin，et al. SmartBoa：Constructing p2p Overlay Network in the Heterogeneous Internet Using Irregular Routing Tables. Proceedings of the 3rd International Workshop on Peer-to-Peer Systems (IPTPS'04)，Feb 2004.

[62] Sergio Marti，Prasanna Ganesan，Hector Garcia-Molina. DHT Routing Using Social Links，3rd

International Workshop on Peer-to-Peer Systems,2004.
[63] Moni Naor. Udi Wieder. Know thy neighbor's neighbor: Better routing for skip-graphs and Small Worlds. Proceedings of the 3rd International Workshop on Peer-to-Peer Systems,Feb 2004.
[64] Gummadi K,Gummadi R,Gribble S,Ratnasamy S,Schenker S,Stoica I. The impact of DHT routing geometry on resilience and proximity. In Proc of SIGCOMM ACM,Karlsruhe,Germany,Sep 2003: 381-394.
[65] Wang S,Xuan D,Zhao W. On Resilience of Structured Peer-to-Peer Systems,in Proceedings of IEEE Global Communications Conference (GLOBECOM 2003)-General Conference,December 2003.
[66] Liben-Nowell David, Balakrishnan Hari, Karger David. Analysis of the Evolution of Peer-to-Peer Systems. ACM Conf. on Principles of Distributed Computing (PODC),Monterey,CA,July 2002.
[67] Nicholas J A, Harvey Michael B Jones, Marvin Theimer, Alec Wolman. Efficient Recovery From Organizational Disconnects in SkipNet. In Proc. IPTPS,Berkeley,CA,USA,February 2003.
[68] Chen Yan,Randy H K,Kubiatowicz J D. Dynamic Replica Placement for Scalable Content Delivery. 1st International Workshop on Peer-to-Peer Systems (IPTPS 2002),March 2002.
[69] Rao Ananth, Karthik Lakshminarayanan, Sonesh Surana, et al. Load Balancing in Structured P2P Systems. In Proc IPTPS,Berkeley,CA,February 2003.
[70] Karger David R,Ruhl Matthias . New Algorithms for Load Balancing in Peer-to-Peer Systems. IRIS Student Workshop (ISW '03) Cambridge,MA,August 2003.
[71] Karger D R, Ruhl M. Simple Efficient Load Balancing Algorithms for Peer-to-Peer Systems. In ACM Symposium on Parallelism in Algorithms and Architectures,June 2004.
[72] Byers John,Considine Jeffrey,Mitzenmacher Michael . Simple Load Balancing for Distributed Hash Tables. In Proceedings of the Second International Workshop on Peer-to-Peer Systems (IPTPS), Berkeley,CA,February 2003.
[73] Xu Zhichen, Zhang Zheng. Building Low-maintenance Expressways for P2P Systems. Internet Systems and Storage Laboratory,HP Laboratories Palo Alto,HPL-2002-41,March 1st,2002.
[74] Xu Zhichen,Tang Chunqiang,Zhang Zheng. Building Topology-Aware Overlays using Global Soft-State. In ICDCS'03,May 2003.
[75] Garcés-Erice L,Ross K W,Biersack E W,et al. Topology-Centric Look-Up Service. In Proceedings of COST264/ACM Fifth International Workshop on Networked Group Communications (NGC), Munich,Germany,September 2003: 58-69.
[76] Freedman Michael J,Mazières David . Sloppy Hashing and Self-Organizing clusters. In Proceedings of the Second International Workshop on Peer-to-Peer Systems (IPTPS),Berkeley,CA,February 2003.
[77] Xu Zhiyong, Min Rui, Hu Yiming. HIERAS: A DHT-Based Hierarchical Peer-to-Peer Routing Algorithm,in Proceedings of the 2003 International Conference on Parallel Processing (ICPP'03), Kaosiung,Taiwan,Oct. 2003: 187-194.
[78] Castro M,Druschel P,Hu Y C,et al. Topology-aware routing in structured Peer-to-Peer overlay networks. MSR-TR-2002-82,September 2002. Available at ftp://ftp. research. microsoft. com/pub/tr/tr-2002-82. pdf.
[79] Castro M, Druschel P, Hu, Y C, et al. Exploiting Network Proximity in Peer-to-Peer overlay networks. Tech. Rep. MSR-TR 2002-82,Microsoft,2002.
[80] Ratnasamy S, Handley M, Karp R, et al. Topologically-aware Overlay construction and server selection. In Proc. INFOCOM'02,2002.
[81] Zhao B Y,Duan Y,et al. Brocade: Landmark routing on overlay networks,Springer,2002.
[82] Saroiu S,Gummadi P K,Gribble S D. A measurement study of Peer-to-Peer le sharing systems. In Proceedings of Multimedia Computing and Networking 2002 (MMCN),San Jose,CA,January 2002.

[83] Zhu Yingwu, Wang Honghao, Hu Yiming. A Super-Peer Based Lookup in Structured Peer-to-Peer Systems. Appears in Proceedings of the 16th International Conference on Parallel and Distributed Computing Systems (PDCS'03). August 2003, Reno, Nevada.

[84] Mizrak A, Cheng Y, Kumar V, et al. Structured Superpeers: Leveraging Heterogeneity to Provide Constant-Time Lookup, Proceedings of the IEEE Workshop on Internet Applications, San Jose, CA, June 2003.

[85] Sean Rhea, Dennis Geels, Timothy Roscoe, et al. Handling Churn in a DHT, In Proceedings of the USENIX Annual Technical Conference, June 2004.

[86] Bamboo. http://bamboo-dht.org.

[87] Christin N, Chuang J. On the cost of participating in a Peer-to-Peer network. Proceedings of the 3rd International Workshop on Peer-to-Peer Systems (IPTPS'04), Feb 2004.

[88] Xu Zhiyong, Min Rui, Hu Yiming. Reducing Maintenance Overhead in DHT Based Peer-to-Peer Algorithms. Linkpings, Sweden, Sept. 2003: 218-219.

[89] Li J, Stribling J, Gil T M, Morris R, et al. Comparing the Performance of Distributed Hash Tables Under Churn. In IPTPS, February 2004.

[90] Mahajan R, Castro M, Rowstron A. Controlling the Cost of Reliability in Peer-to-Peer Overlays. In IPTPS'03, Feb 2003.

[91] Rodrigo Rodrigues, Charles Blake. When Multi-Hop Peer-to-Peer Routing Matters. Proceedings of the 3rd International Workshop on Peer-to-Peer Systems (IPTPS'04), Feb 2004.

[92] Prasanna Ganesan, Krishna Gummadi, Hector Garcia-Molina. Canon in G Major: Designing DHTs with Hierarchical Structure, 24th International Conference on Distributed Computing Systems (ICDCS'04), Hachioji, Tokyo, 2004: 24-26.

[93] Bujor Silaghi, Bobby Bhattacharjee, Pete Keleher. Query Routing in the TerraDir Distributed Directory. In Victor Firoiu and Zhi-Li Zhang, editors, Proceedings of the SPIE ITCOM 2002, SPIE, Boston, MA, August 2002, 4868: 299-309.

[94] Garcés-Erice L, EBiersack W, Felber P A, et al. Hierarchical Peer-to-peer Systems. Proceedings of ACM/IFIP International Conference on Parallel and Distributed Computing (Euro-Par), 2003.

[95] Jeffrey Considine. Cluster-based Optimizations for Distributed Hash Tables, 2002. http://cs-www.bu.edu/ftp/fs/techreports/pdf/2002-031-DHT-cluster-optimization.pdf.

[96] Ganesan Prasanna, Sun Qixiang, Garcia-Molina Hector. YAPPERS: A Peer-to-Peer Lookup Service over Arbitrary Topology. In Proceedings of the Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM), San Francisco, CA, March 2003.

[97] Kempe David, Jon M Kleinberg, Alan J Demers. Spatial Gossip and Resource Location Protocols. In Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing (STOC), pages 163-172, Crete, Greece, July 2001.

[98] Loo B T, Ryan Huebsch, Ion Stoica, et al. The Case for a Hybrid P2P Search Infrastructure. Proceedings of the 3rd International Workshop on Peer-to-Peer Systems (IPTPS'04), Feb 2004.

[99] Gupta I, Birman K, Linga P, et al. Kelips *: building an efficient and stable P2P DHT through increased memory and background overhead. In Proceedings of the Second International Workshop on Peer-to-Peer Systems (IPTPS), Berkeley, CA, February 2003.

[100] Yang B, Garcia-Molina H. Improving Search in Peer-to-Peer Systems. In Proc of the 2nd International Conference on Distributed Computing Systems, Vienna, Austria, July 2002.

[101] Matthew Harren, Joseph M Hellerstein et al. Complex Queries in DHT-Based Peer-to-Peer Networks. Lecture Notes in Computer Science, Cambridge, MA, March 2002, 2429: 242-250.

[102] Reynolds P, Vahdat A. Efficient Peer-to-Peer Keyword Searching. In Unpublished Manuscript,

June 2002.

[103] Bloom B H. Space/Time Tradeoffs in Hash Coding with Allowable Errors. Communications of the ACM,July 1970,13(7): 422-426.

[104] Bobby Bhattacharjee, Sudarshan Chawathe, Vijay Gopalakrishnan, et al. Efficient Peer-to-Peer Searches Using Resultcaching. In The 2nd International Workshop on Peer-to-Peer Systems (IPTPS'03), February 2003.

[105] Bryce Wilcox-O'Hearn. Experiences deploying a large-scale emergent network. In Proceedings of the First International Workshop on Peer-to-Peer Systems (IPTPS' 02), Cambridge, MA, March 2002.

[106] Sun Qixiang, Garcia-Molina, Hector. Partial Lookup Services (Extended Version), Technical Report. Stanford University,2002.

[107] Ozgur D Sahin,Abhishek Gupta,Divyakant Agrawal,et al. A Peer-to-Peer Framework for Caching Range Queries,ICDE 2004.

[108] Sahin O,Gupta A,Agrawal D,et al. Query processing over Peer-to-Peer data sharing systems. Technical Report UCSB/CSD-2002-28,University of California at Santa Barbara,2002.

[109] Gupta A,Agrawal D,Abbadi A E. Approximate Range Selection Queries In Peer-to-Peer Systems. In Proceedings of the First Biennial Conference on Innovative Data Systems Research (CIDR), Asilomar,CA,January 2003: 141-151.

[110] Nathan Linial,Ori Sasson. Non-Expansive Hashing. In Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing (STOC),Philadelphia,PA,May 1996: 509-518.

[111] Joseph S. P2P MetaData Search Layers. Second International Workshop on Agents and Peer-to-Peer Computing AP2PC 2003.

[112] Zhang Zheng, Shi Shuming, Zhu Jing . SOMO: self-organized metadata overlay for resource management in P2P DHT. In Proceedings of the Second International Workshop on Peer-to-Peer Systems (IPTPS),Berkeley,CA,February 2003.

[113] Crespo A, Garcia-Molina H. Routing Indices For Peer-to-Peer Systems. Proceedings of the International Conference on Distributed Computing Systems (ICDCS),July 2002.

[114] Jordan Ritter. Why Gnutella Can't Scale. No, Really. 2001, www. darkridge. com/～jpr5/doc/gnutella. ntml.

[115] Fast Track. http://www. fastrack. nu.

[116] Harvey N J,Jones M B,Saroiu S ,et al. Skipnet: Ascalable Overlay Network with Practical locality properties. In Proc of USITS,2003: 113-126.

CHAPTER 10

第10章　对等网络的安全问题

摘要：对等网络的安全问题日益得到强调和重视。本章针对对等网络的节点的自私行为和恶意行为，介绍了当前对等网络的激励机制、信任机制以及文件安全机制的研究工作。

关键字：P2P、激励、信任、文件真实性、文件污染。

10.1　研究背景

随着现有P2P系统的广泛、深入地应用，逐渐暴露出P2P系统的一些安全问题，其中一个主要原因是存在一些诸如节点自主行为引起的不可靠的服务质量、网络攻击等。P2P系统具有的匿名性、高度的开放性以及加入系统的用户节点类型、目的、利益空间差异性大等因素致使节点行为也不尽相同，按照节点自主行为的不同将这种情况分为以下两种：

（1）自私行为。如"搭便车"（free-rider）和"公用地悲剧"（common tragedy）。其中"搭便车"指节点只使用其他节点提供的资源或服务，而不共享自己的资源。以Gnutella文件共享系统为例，高达70%的节点是"搭便车"[33]。公用地悲剧指网络资源作为一种非排他的公共资源，被大多数P2P节点无节制的使用。节点之间的不合作及利己的自私行为严重影响了P2P服务的可用性。

（2）恶意行为。P2P网络中还存在着为数不少的蓄意提供不可靠的服务质量以及欺诈行为的节点，这些节点并不关心资源访问而仅仅为了散布无效的或有害的内容，如虚假的文件，病毒和木马等，它们对所有的查询都做出积极的响应（无论是请求下载文件还是请求推荐信息）。恶意节点的存在严重影响P2P系统的性能，使得P2P系统的可用性得到极大的破坏。

本章通过激励机制和信任机制、文件安全机制对节点的行为进行约束，引导P2P节点更倾向于成为一个"好"行为的节点，构建一个良性安全的对等网络。

10.2　激励机制

P2P系统尽管具有系统容错性高、容量大、可扩展性好等优点，但是缺点同样明显。由于P2P网络的分布式特性，单个节点具有更大的自由度，服务

质量因此无法得到保证：节点动态性使得服务不稳定：节点的自私性导致只有少部分节点愿意提供服务；同时少部分恶意节点更容易对其他节点发起攻击，这些攻击包括针对文件的攻击和针对路由的攻击[2]。系统测量表明[3]：P2P 网络中大量用户（约 70%）只获取服务，不提供服务，这导致了 P2P 网络较为严重的不公平性问题。有的节点只获取服务，不提供服务[3]；还有一部分节点通过欺诈手段增大自身利益，针对文件进行攻击。例如随意伪造文件，随意终止服务等[4]；甚至有的节点之间互相勾结联合进行攻击（sybil attack）[1]。在传统网络中，由于服务总是由中心服务节点提供，公平性问题和服务的安全性可靠性问题并不突出，但是在 P2P 中，我们需要激励节点提供安全可靠的服务。

10.2.1　搭便车问题

P2P 网络资源的丰富依赖于用户对自己可用资源的共享。但是，真正在网络上创造或提供内容的人还是少数的，据统计 Gnutella 的用户中仅仅有 2%向其他用户提供了内容，即使在比较活跃的 Usene 七张贴文章的用户也仅占所有用户的 7%。虽然 P2P 模式中有着传输速度大和共享性质的种种优点，但现实中往往出现人们只想得到别人的共享文件却没有把自己的有用资源共享出来的问题，由此产生了 P2P 网络应用中的一个广泛存在的消极现象：搭便车。搭便车定义为一个自私的个体有意识地拒绝为某个群体的共同利益自愿地贡献。Saroiu 等在 2002 年发表的论文[5]中指出根据他们最新对 Napster 和 Gnutella 的实验结果表明：大多数用户只作为消费者而不对系统贡献自己的资源，50%“种子”在线时间不超过 1 小时，多数的用户都是搭便车者，对系统贡献很少，如 Gnutella 系统中有 25%的用户根本不共享任何资源。“种子”的短缺意味着系统中部分有价值的资源长时间得不到利用，搭便车现象的增加使得整个系统失去 P2P 分布式共享资源的精神，这一现象的蔓延，将导致 P2P 退化成传统的 C/S 系统[6]。由此引发了许多关于 P2P 激励机制方面的理论研究课题和商业计划，提出 P2P 激励模型，向系统贡献资源的用户能从网络中得到相应的回报，激励用户共享资源以消除搭便车现象。

10.2.2　激励机制

在 P2P 网络中，尤其是纯 P2P 模式的网络中，并没有一个中心节点或是权威对合作进行监督。每个节点都是理性的，它们追求的是自身效用的最大化。如果与其他节点合作，提供资源或服务，将会消耗节点的资源，并降低节点的性能。尽管理性的节点可以预见缺乏合作会导致 P2P 网络的总体性能的下降，但由于合作所带来的好处并不直接，因此节点合作的动力不大。有研究人员对一个具有 35 352 台主机的实际运行的 Gnutella 网络进行了 24 小时的监控发现：在 Gnutella 网络中搭便车现象盛行，近 70%用户是搭便车者，即它们未被共享任何文件；近 50%的回应来自仅 1%的用户。研究人员认为：P2P 网络需要自发的合作，但这对于匿名的大规模系统很难做到；理性用户关心效用，并会影响其是否选择合作。

一个缺乏合作的 P2P 网络中，搭便车现象非常普遍，很多节点之下载资源或索取服务，从来不提供资源或服务。更有甚者，有的节点提供假的资源或服务，这种现象在 eMule 网络中也屡见不鲜。有些节点故意限制或谎报网速，以达到少提供或不提供资源或服务的目

的。所有的这种种行为,导致了公用地悲剧[7]:网络资源集中化,造成网络拥堵。另一方面也导致网络资源同质化,节点不愿意主动引入新的资源。最终使得整个 P2P 网络性能急剧下降,所有参与其中的节点的利益都受到不同程度的损害。

上述问题的根源在于,目前的 P2P 网络,无论是采用何种架构或模式,都基于一个假设,那就是每个参与的节点都能善意地,最大化地提供网络资源。这个假设忽略了一个事实,那就是每个节点后面都是具体的人,节点的行为实际上就是个人的行为。因此,可以通过研究个人的行为来解决这个问题。作为组织行为学的重要组成部分的激励机制,在这里可以发挥巨大的作用。

1. 面临的问题

尽管传统的激励机制在实际应用中已经展现出它的巨大威力,然而 P2P 网络具有其特殊性,因而不能照搬传统的激励机制,而是应该从 P2P 的实际情况出发,制定出适用于 P2P 网络的激励机制。在制定之前,需要对 P2P 网络中激励机制可能面临的问题[8]进行研究。

传统的组织有大有小,但总的来说,一个组织的员工并不会太多。以全球最大的软件公司——微软为例,其业务已经遍及全球 90 多个国家和地区,其全球员工总数将近 60 000 人,它可以算是一个巨型的组织了。而根据统计[5],在一个像 Gnutella 和 KaZaa[9]这样的文件共享系统中,光是其并发的用户数完全可能超过 100 000。由此可见,不同于传统组织,P2P 网络可以拥有非常多的成员(节点)数量。

在传统的组织中,尽管也存在着员工离职或新员工入职,从而存在一定的周转率,一般来讲,这个周转率并不会太高(如果有组织有很高的周转率的话,工作效率的降低往往是因为员工交接及新员工需要时间适应新环境和新工作引起的,这并非激励机制所能解决的)。然而,根据研究人员的统计,在像 Gnutella 和 KaZaa 这样的文件共享系统中,节点从加入网络到离开网络的时间间隔,即"生存时间",其平均值基本上是以分钟计算的。由此可见,P2P 网络具有很高的周转率。

传统的组织中,员工往往以部门或项目为单位进行组织管理。在同一个单位中的员工,往往具有相同或类似的兴趣,这为激励机制的应用带来了很大的方便。而在 P2P 网络中,由于节点都是对等的,因而并不存在这样的单位。而且,节点之间的兴趣也是不对等的,比如 A 可能对 B 的资源或服务感兴趣,而 B 却对 C 的资源或服务感兴趣。P2P 网络这种兴趣不对等的特征对于传统的激励机制来讲,是一个巨大的挑战。

在 P2P 网络中,多个节点可能串通作案,以提高自身的信誉。这种情况在一些采用了基于信誉的激励机制的 P2P 网络中非常普遍。在传统的组织中,尽管也存在类似的问题,但一方面,传统的激励机制中信誉不是主要的因素,因而 P2P 文件共享系统中激励机制的研究这种行为对传统激励机制的影响比较小;另一方面,在 P2P 网络中的这种行为更难察觉,具有更大的危害性。

P2P 网络还带来了两个在传统的组织中不存在的问题。一个是零成本身份切换,在 P2P 网络中,由于节点都是匿名的,因而节点可以任意地更改自己的身份。另一个是行为背叛的问题,即一个在历史上行为良好的节点,突然更改其行为,变成一个搭便车者,或作出其他危害 P2P 网络的行为。

2. 激励模型

在 P2P 网络中,激励模型可以大体被分为两类:

(1) 货币支付模型,节点间每次文件的上传和下载都需要明确的货币支付。

(2) 非货币模型(也被称为软激励模型),不使用明确的货币支付而采用其他的方法提供激励。

1) 货币支付模型

(1) 真实货币支付。用户在网络中出售自己的资源,利用了用户资源的网格组织通过银行向该用户支付真实货币,付费在网络外进行。这种方式广泛地在网格计算的经济模型中,目前的一些网格系统如 Globus、Legion 等已经提供了大量的、成熟的、可重用的中间件,例如资源协同分配服务 DUROC、认证和安全服务 GSI 等。

(2) MicroPayment。在这种模型下,所使用的货币并不能在网络以为兑换现实货币,可以称为虚拟货币。每个用户下载资源之前必须向服务提供节点支付相应价格的虚拟货币。为网络中其他节点提供服务可以得到相应的虚拟货币,所以为了能够持续的得到网络资源,节点必须不断地以自己的服务换回足够多的虚拟货币才行。其中货币的交易必须具备 ACID(Atomicity 原子性,Consistency 连贯性、Isolation 孤立性、Durability 耐久性)。现实中的系统包括 MojoNation[16]、Maze 等。

2) 软激励模型

(1) 实物交换模型。用户总是只有共享了文件,对系统做出了贡献才能从别人那里下载文件。上传与下载的 速率同样受到控制,即用户使用一定的速率下载文件,也必须要提供相应的上传速率。使用这样的机制的目前有 eDonkey 和 BitTorrent。

(2) 区分服务质量模型。向系统贡献越多,则可以得到更好的服务,比如优先的访问权,更为稳定的传输,更高速的下载,更大的存储空间等。例如在文献[10]一文中所提出的基于节点服务情况和使用情况矩阵特征向量来决定是否向请求服务提供节点提供其所要求的服务。

对于如何在 P2P 上建立激励模型,采用什么样的激励机制及算法,众多研究者进行了大量的研究和实验。目前在 P2P 网络中引入激励机制的主要方法有:

① 重新设计分布式的资源定位协议来实现网络激励,如对于网络贡献大的节点可以把资源搜索报文传输到更多的节点,进而增大节点资源的发现概率。

② 在已有的协议上增加特殊的激励算法,如请求的接纳控制,服务质量的区分等来实现网络激励。

现在大量的研究工作集中在现有路由协议上增加特殊激励算法来实现激励这方面,主要是因为现有的资源定位协议,不论是使用无结构洪泛还是使用有结构分布式 HASH 函数来定位资源的协议都基本已经成熟,并且有大量的应用,抛弃现有的大量系统而重新开发基于新协议的应用代价过大。采用在现有协议上增加激励的方法,不仅理论上更容易实现,而且解决了现实中大量已存在系统的缺陷,有更加现实的意义。

10.2.3　激励机制研究现状

1. 基于微支付和虚拟货币的机制

斯坦福大学的 Phillipe Golle 等人首先提出了使用微支付和虚拟货币的方法来解决

P2P 共享网络中的共享激励问题[11]。这个机制主要思想是，服务器记录每个注册用户文件下载数量 u 和文件上传数量 v，每次用户间传递文件时，服务器将提供文件下载的用户的文件上传数量 v 加 1，并将下载文件的用户的文件下载数量 d 加 1。每隔一段时间，服务器为每个用户计算应支付的金额 $C=f(u,v)$。这里的函数 f 根据一定的算法将用户的下载上传差额换算成用户应该支付的金额。函数 f 一般采用线性函数，以便使得整个网络中的微支付总额为 0(收支平衡)。

为了增加该机制的灵活性，该机制又引入被称为“点数”的虚拟货币。服务器不再记录用户的文件下载数量和文件上传数量，而是记录用户的点数。每次用户间传递文件时，服务器增加提供文件下载的用户的点数，同时减少下载文件的用户的点数。用户可以通过提供文件下载来获得点数，也可以直接使用现实货币购买点数。在使用现实货币购买点数的情况下，由于交易额一般很小，直接通过银行支付费用相对较高，而且比较麻烦，因此需要有第三方来向用户销售点数，再由其与银行结算，这个第三方被称为经纪人(broker)。

这样一来，经纪人节点的稳定性在这种机制中就至关重要了。因此，PPay[12]引入并实现了浮动的、自管理的货币的概念，以最大限度地减少经纪人的介入。浮动的货币允许数字货币(即虚拟货币)在没有经纪人参与的情况下，也能从一个节点“浮”到另一个节点。至于说自管理，是指在 PPay 中，由拥有数字货币的节点自身来负责数字货币的安全，只有在购买数字硬币或兑换成现实货币时才需要经纪人的介入。

Fileteller[13]是一个网络文件存储系统，它也采用微支付方式来对网络中的用户进行激励。

在支付系统中，资源或服务以商品的形式存在，消费者要购买使用权，而提供者可以获得报酬。P2P 系统中，节点通过向其他节点提供资源或服务来获得报酬，并用获得的报酬去购买自己所需的资源或服务。由于 P2P 系统中的资源常常是比较廉价的(例如空闲的带宽或 CPU 时间等)，所以在 P2P 中采用的多是微支付(micropayment)系统，每笔交易的价值较小，对安全性等方面的要求也相对较低。

学术界研究文献中提出的比较典型的 P2P 支付系统有 PPay[12]、KARMA[14]、PeerMint[15]等。实用中的 P2P 微支付系统实例有 Popular Power 和 MojoNation[16]，前者付少量的报酬购买别人的空闲 CPU 时间用以构造 P2P 分布计算网络，再出售给需要的用户，后者是一个 P2P 在线支付系统。

2. 基于配额的机制

基于配额的机制的主要思想是为每个节点设定一个配额，节点在一个时间段内从 P2P 网络中的下载总量不得超过这个配额。

CFS[17]是一个只读的 P2P 网络存储系统，为了减少攻击，提高系统的安全性，它采用了存储配额机制，规定每个节点只能访问存储系统总量的一个很小的比例，例如 0.1%。

CFS 引入配额的动机并非是为了激励，而是为了提高系统的稳定性和安全，而 FARSITE[18]和 Pastiche[19]则将配额用于激励机制。这两种激励机制都是通过限制节点能够从 P2P 网络中或取的资源与其对 P2P 网络的贡献相当，从而激励用户贡献资源。

Samsara[20]则采用了一种比较特别的策略：每个节点在向其他节点请求存储空间之前，必须允诺对方能够使用本节点上相同大小的空间，对于那些不遵守规则的节点，Samsara 会以一定的概率进行惩罚。

3. 基于信誉的机制

信任是指一个节点基于个体体验对另一个节点在系统中可信度方面的一个评价，而信誉则是指一个节点通过合作的方式，基于自己或者其他节点的一些信息来获得其他节点在系统中的可信度方面的一个评价。

P2P 系统中信任和信誉关系的基本思想是用户间完成交易后，可以对这次交易进行评价，从而给对方一个评价。用户间可以通过这些相互间直接的评价来建立对对方直接的信任关系。同时，这种直接的信任关系可以通过某种信任传播算法来描述用户在系统中的主观或者客观的信誉值。

信誉模型主要分为主观的信誉模型和客观的信誉模型。主观的信誉模型是指每个节点都建立并维护一个信誉表，这个表中存储了节点对每个有过交往的节点的信誉的评价。而客观的信誉模型中用户不用单独维护信誉表，而是在整个 P2P 网络中，所有节点以某种机制共同维护一个全局信作表，这个表中存储了所有的刀节点的全局信誉值。

SLIC[21]是一个典型的主观的信誉模型。它的基本思想是：每个节点统计各个邻居节点对本节点发出的请求的响应情况，然后根据响应情况给其打分，响应越好，得分越高，这个打分每隔一段时间进行一次，打分的结果作为接下来的时间段中本节点对邻居节点的请求进行响应的依据。每个节点既要响应邻居节点的请求，又要发出自己的请求，而且每个节点的能力是有限的，因此它必须要在响应请求和发送自己的请求之间找到一个平衡点，使得在为邻居借点服务的同时不影响自己从 P2P 得到服务的质量。

主观的信誉模型实现比较简单，只要发生交互的两个节点的参与，但这意味着节点间对其他节点的评价不尽相同。对于那些为邻居节点提供过良好服务的节点，当它们与新的节点交互时，其良好的历史记录并不能为其带来好处，这样对它们来讲并不公平，一定程度上会降低其积极性。而且容易造成节点的功利性：需要时便向邻居节点提供服务以得到好的服务，不需要时便拒绝提供服务。

客观的信誉模型实现起来要比主管的信誉模型复杂，而且容易受到攻击和欺骗。然而，相比于主观的信誉模型，它帮助形成了更公平的环境，节点即使当前不需要获得服务，它也可以提供良好的服务，提高自己的信誉值，以便将来或遇到新节点时能够得到更好的服务。

4. 基于 TFT 的机制

TFT(Tit-For-Tat)是一种非常简单的合作策略：第一次交易中总是选择合作，之后每次交易采用的策略与对方在上次交易中所采用的策略(合作或欺骗)相同。正如 R. Axelrod 在 *The Evolution of Cooperation* 一文[22]中所指出的那样，TFT 是自我主义者构成的无中心交易环境中最好的合作策略。目前最受欢迎的 P2P 文件共享和内容分发系统 BitTorrent (BT)就使用了这种动机机制，并在改善公平性方面取得了良好的效果[23]。

在 BT 中，TFT 策略通过“阻塞(choking)”算法来实现。在一个 Torrent(同一个文件的下载群)中，每个节点周期性地计算着其所有连接上的下载速度，并根据这些信息，选择其中速度最快的 w(w 默认为 7)个连接提供上载，其他连接除了一个由“乐观疏通(optimistic unchoking)”选定的之外，全部予以阻塞(choke)。乐观疏通的目的是为了有机会找到一个更好的连接。如果该连接上的下载速度优于某个目前正在提供上载的连接，则这个新连接接替其在 w 个连接中的位置；否则，下一轮将以 Round-Robin 方式选择另一个乐观疏通的

连接[23]。

BT 的激励机制可以使尽量多的有效连接处于双向传输的最佳利用状态，并且，为了获得更快的下载速度，每个节点也会有足够的动机在下载的同时也为其他节点提供上载服务，从而有效地减少了搭便车现象的发生。

5. 基于相似性的机制

2001 年，R. L. Riolo 发表在 *Nature* 上的一篇论文[24]提出了另一种合作机制。该文认为除了血缘关系和互惠关系（直接或间接）之外，合作还可能会在“相似”的实体之间产生。实体的特征可以用 Tags 来表示，这些 Tags 被用来进行相似性的度量。文献[25]中提出了一种将 Tags 模型应用于 P2P 合作问题的框架。

10.3 信任机制

P2P 网络是一种典型的“开放式”网络环境。系统中的用户并非来自同一个利益团体，任何人只要愿意都可以自由地加入和退出 P2P 网络，自由地参与资源共享和交换。参与节点的身份对等，所有节点既可以是资源或服务的消费者，也可以是提供者。用户具有很强的自主性，因此提供的服务质量不可能像传统的 C/S 或 B/S 模式那样可靠，它们可以随意终止服务，甚至可能存在欺诈行为。另一方面，P2P 中的交互模式是点对点的，个体间的交互和协作是系统的主要业务，个体行为和节点之间的对等交互不被第三方直接介入或监控。

P2P 模式的开放性、对等性、自主性和无监督性是它的主要特征，也是它在很多领域取得巨大成功的重要原因。但这种个人为公众提供资源同时又享受公共资源，而节点行为无约束的工作模式，使 P2P 网络中“信任”极度缺乏，交互双方很难判断对方的可信程度，交互对象的选择具有很大的盲目性，服务质量和安全也很难得到保证。

P2P 网络中的大量不可靠服务都是由于信任缺失而引起的。例如在众多文件共享 P2P 网络中，大量的文件是伪造的、未经授权的或被篡改过的，甚至是带有病毒的。而在类似于 eBay 和淘宝网这样的电子商务类 P2P 系统（广义）中，这种不可靠服务和欺诈行为给用户带来的影响则更为严重。

10.3.1 信任概念

信任是一个多学科的概念，描述了在特定的情境下，一个个体 A 在可能产生不利后果的情况下（包括风险因素），愿意相信另一个个体 B 具有某种能力或能够完成某项任务的主观信念[26]，或个体 A 根据自己的经验或同时参考其他个体 C、D 等的推荐信息而得出的被信任方 B 的可信赖程度。Luhmann 于 1979 年从社会学的角度来描述信任，将其定义为减少社会复杂性的方法[27]。而这种复杂性是由具有不同理解力和目的的个体的交互引起的。该定义由于其社会学的本质更适合基于信誉的系统。另一个广为接受的定义是计算机科学家 Gambeta 于 1990 年给出的[28]。信任被定义为个体评估另一个体或集体将执行某一特定行为的特定主观可能性等级，评估发生在个体能够观察到该特定行为之前（或该特定行为独立于个体能够观察到该行为的能力）且该特定行为会影响评估者自身的行为。

Gambeta 认为信任不是一个门限点，而应该是一个概率分布的概念，可以用介于完全

不信任(用 0 表示)和完全信任(用 1 表示)之间的值来表示,且以不确定性为中点(用 0.5 表示)。该定义引入了从信任方的角度认识到的被信任方的可靠性概念。最近的关于信任的概念是由 Grandison 和 Sloman 提出的[29],他们将信任定义为对某一个体在特定的情境下,独立、安全且可靠的完成任务的能力的坚固信念。Chen R[30]认为:信任是大多数人际关系的核心,信任的因素因人而异,每一个人都有它自己的意见,因此信任的本质是分布的。

与信任紧密联系的概念是信誉,Abdul-Rehman 将信誉定义为基于观察到的个体过去行为或过去行为的信息而对个体行为的期望[31]。可以看出,信誉强调的是一个集体对某一个体(或群体)的综合的可信赖度,而信任更多强调的是信任个体对被信任方的主观信赖。

10.3.2　信任模型

信任模型(trust model)是指建立和管理信任关系的框架。信任模型分为两种基本类型:层次信任模型、网状信任模型。层次信任模型是较为简单,并广泛使用的信任模型(如 X.509),其优点是结构简单,易于管理和实现,缺点是信任关系必须通过根来实现,层次模型适合孤立的、层状的封闭环境。网状信任模型中,每一个节点都可以作为可信任根,节点间的信任路径可以构成一个网络,如 PGP[32],网状信任模型的优点是更接近于人类社会的信任关系,信任关系易于构建,且不依赖于任何权威中心。

10.3.3　信任模型的研究现状

信任模型是建立和管理信任关系的框架。目前,P2P 信任研究主要涉及信任量化、信任评价以及信任的计算、存储、传递机制的研究。相对于传统的安全技术,信任模型更像是一个不十分"严格"的安全技术。它不像一般的鉴别、认证等技术,有一个明确的接受或者拒绝的标准,而是更具有主观性。跟传统的安全技术相比,它建立了一个类似人类社会的信任评价和信誉传递机制,能更好地处理安全中的信任问题,一旦和已有的安全技术结合起来,就能对解决对等网中的安全问题提供较好的解决方案。

在分布式系统中,建立不同网络节点间的信任关系是建立系统安全的一个基础。P2P 信任机制主要是用来解决如何选择可信赖的 Peer 的问题的。信任问题在 C/S 时代也是存在的,但是在 C/S 时代构建信任机制要容易得多,因为 Server 处于中心位置,可以方便地收集 Client 的各项信息。但是由于对等网络的特点,信任模型是 P2P 网络应用中一个十分难以解决的问题。到目前为止,P2P 信任机制已经成为一个活跃的研究课题,尽管目前在实际应用中还没有出现比较成功的通用解决方案,但国内外研究者在信任研究领域开展了许多开创性的研究工作,一些具有代表性的研究成果提出了不少值得借鉴的思路和方法。

1. 集中式信任模型

集中式信任系统主要应用于电子商务领域,在实际应用中大都采用基于 PKI 的信任模型,简单地采用给参与者评价打分的方法来描述信誉度,采用简单的数值计算方式来实现信任的聚合。

在集中信任系统中,存在少数中心实体负责收集网络参与实体的历史交易记录信息,然后再把所有实体的信誉评分的结果公布出来。在下次的实体交易前,请求实体即可以通过参考备选服务实体的最新信誉信息来加以选择。这样拥有良好信誉的服务实体会获得更多

的提供服务机会和回报,与此同时诚实可信的实体也会获得更高的信誉,从而抑制网络中的不良行为,最终会促进网络的良性发展。

集中式信任系统中,实体 A 和 B 在历史交易记录的基础上,在信任系统的支撑下相互选择对方,因为双方均认为对方的行为最可靠,从而开始一次新的交易交互。在每次交易结束后,实体之间会对对方在交易中的行为给出评价,信誉中心会不断地收集每个实体的评价信息并更新每个实体的信誉信息,更新后的实体信誉信息会在信誉中心公布。

在这类系统中,中心实体负责整个网络的监督,定期通告违规的实体,中心实体的合法性通过 CA 颁发的证书加以保证。这类系统往往是中心依赖的,具有可扩展性、单点失效等问题。这类实际系统的实例有 eBay、eDonkey 等。

2. 基于局部推荐的分布式信任模型

在这类系统中,节点通过询问有限的其他节点以获取某个节点的信誉度,一般采取简单的局部广播的手段,其获取的节点信誉度是局部的。如 Comelli[33] 对 Gnutella 的改进建议就是采用这种方法。

在局部信任模型中,许多研究者认为信任是主观的,对同一实体的可信度,不同的观察者可能会得出不同的判断。例如,在 P2P 环境中,不同的参与者可能会采用不同的方法来评价其他参与者的性能,这反映了不同用户对行为的不同理解[34](信任的上下文相关性特征)。在基于推荐的局部信任模型中,参与者通过询问其他有限的参与者(推荐者)来获得某个参与者的信任度信息,从而形成自身的信任观点。目前基于局部推荐的分布式信任模型的主要研究工作有:

(1) 基于概率的局部信任模型。在信任研究领域,部分研究者认为借助于概率方法可以描述主观信任,提出了多种基于概率的信任模型。

文献[35]采用 Bayesian 公式对实体的信任相关经验进行了建模,提出了对实体间信任度进行定量研究的 Beth 模型。除 Beth 模型外,文献[34]针对 P2P 文件共享系统,提出了一个基于 Bayesian 网络的信任管理模型。该模型认为信任来自于参与者的直接经验,信誉则基于其他参与者的推荐,信任都具有上下文相关性、多面性、动态性等特点。

(2) 基于主观逻辑的信任模型。Josang 模型中将行为的结果(成功或失败)作为经验,根据二项事件(binary event)后验概率服从 Beta 分布的思想,提出了基于主观逻辑(subjective logic)的信任度评估模型来解决信任的推导和综合计算[36]。

(3) Abdul-Rahman 模型。与基于概率的局部信任模型不同,Abdul-Rahman 等[37]认为,尽管从直观上看,信任度可表示为某种概率的度量,但问题在于概率值只有以定义明确的可重复实验为基础才有意义,因而不适于处理日常的实际经验。并且,基于概率的模型仅仅考虑了观察本身,没有考虑观察者。此外,概率本质上是传递的,而信任只具有弱传递性。

Abdul-Rahman 模型将信任划分为 4 级,分别累计不同级别的交易经验,并以此为基础进行信任的评价和推理。Abdul-Rahman 模型的不足之处在于其信任的表示和推理方法比较复杂,缺乏直观意义。

(4) PeerTrust 模型。Xiong Li[4]给出了一个适用于 P2P 电子社区的局部信任模型,节点的可信度是对以往该节点向其他节点提供服务的水平的综合评价。模型考虑全面,引入了节点对交互的反馈,反馈的可信度,节点参与交互的次数,交互的属性和节点所在社区五个因素度量节点的可信程度。其信任值的计算公式为

$$T(u)=\alpha*\sum_{i=1}^{I(u,v)}S(u,i)*\mathrm{Cr}(p(u,i))*\mathrm{TF}(u,i)+\beta*\mathrm{CF}(u)$$

相关参数的定义为：$T(u)$是节点 u 的信任值，$\mathrm{Cr}(v)$是节点 v 的推荐意见的可信程度，$\mathrm{TF}(u,i)$是节点 u 第 i 次交互的属性，$\mathrm{CF}(u)$是节点 u 所在社区的属性，$I(u,v)$是节点 u 和 v 之间交互的总数，$p(u,i)$是节点 u 第 i 次交互的对象，$S(u,i)$是归一化的节点 u 的第 i 次交互后 $p(u,i)$对它的信任评分，α 是与 u 交互过的节点对 u 的综合评价的权重，β 是社区对评估的影响所占的权重。

该模型对实体得到的推荐信任进行统计和分类计算得到实体的信任度，模型认为需要识别欺骗行为和对欺骗者进行惩罚，却没有提出具体的方法和机制。

(5) 其他局部信任模型。除上述局部信任模型外，其他研究者还提出了基于轮询投票[33,38]的信任模型。其中，Damiani 等基于 P2P 文件共享协议 Gnutella 提出了 Damiani 模型[33,38]。该模型以 Gnutella 协议的资源搜索与响应消息 Query 和 QueryHit 为基础，提出了轮询协议 XRep，参与者在下载资源之前先请求其他实体对某目标实体进行投票，投票的消息采用公钥技术实现签名与加密，在对投票结果进行聚集分析和 challenge/response 检查，排除可疑投票之后，根据其结果确定是否访问该目标实体。在实际与目标实体进行交易之后，还要更新目标实体及投票者的可信度。

3. 基于全局推荐的分布式信任模型

为获取全局的实体可信度，该类模型通过实体间相互满意度的迭代，从而获取实体全局的信誉度。Stanford 的 EigenRep[39]是目前已知的一种典型的全局信任模型，该信任模型在信任问题研究领域具有重要的指导意义，成为大部分研究工作的参考标准。EigenRep 的核心思想是：依据一个节点提供的成功交易的次数与失败交易的次数来计算该节点的信誉值。当节点 i 需要了解任意节点 k 的全局可信度时，首先从 k 的交易伙伴中(曾经与 k 发生过交易的节点 i)来获知节点 k 的可信度信息，然后根据这些交易伙伴自身的局部可信度(从 i 的主观判断角度来看)综合计算出 k 的全局可信度。计算公式如下

$$T_k=\sum_j C_{ij}\times C_{jk}$$

式中：T_k 为节点 k 全局的可信度，对于任意节点 i、j、C_{ij} 为节点 i 对节点 j 的局部信任度，计算公式如下

$$C_{ij}=(\mathrm{Sat}_{ij}-U_n\mathrm{Sat}_{ij})\Big/\sum_j(\mathrm{Sat}_{ij}-U_n\mathrm{Sat}_{ij})$$

式中：Sat_{ij} 和 $U_n\mathrm{Sat}_{ij}$ 分别为节点 i 对 j 在历史交易中积累的满意次数和不满意次数。

文献[52]分析了 EigenRep 模型存在的不足之处，如缺乏迭代收敛性保证、没有考虑惩罚因素和网络性能开销等，提出了基于推荐的 P2P 环境下的 Trust 模型，进行了相关分析并给出了分布式计算协议。

4. 基于组群的 P2P 信任模型

基于组群的 P2P 信任模型[53,54]以信誉为基础，通过将节点组织成组群来实现 P2P 系统安全控制。在文献[53]中，通过计算一个双层信誉，以此为依据最终选择一个节点进行交易，从而提高网络交易的质量和安全性。在该模型中，主要是通过逻辑上的一个节点组来实现双层信誉的(节点所在组的信誉和该节点本身的信誉)。模型规定每个节点在同一时间至

多只能属于一个组。通过该节点所在组的信誉和该节点本身的信誉,可以判定这个节点是善意的或是恶意的,是否可以信任。一个节点的信誉可以通过该节点所要执行动作善恶可能性的大小来评定。并且随着不同节点之间的交互动态调整。如果一个节点滥用系统中的资源或者有自私的行为,该节点将受到降低信誉的惩罚。同样,如果有节点试图通过系统的匿名性伪装自己来攻击其他用户或更改其他用户的路由信息,也将受到惩罚。当其再次试图寻求协作时,系统将会提示这些节点是恶意的。另一方面,系统也将会给可靠诚实的节点一个好的信誉。当一个节点拥有较高的信誉时,该节点被认为是可信任的、并且可以获得所在组的支持。一个节点的可靠性越高,希望与其交互的节点就会越多。模型采用了双层的信誉模式,每一个节点行为的好坏同时还会影响所在组的信誉,这样可以激励组内的所有成员相互监督。在这种"监督"的压力之下,每一个节点都必须尽量做到最好,因为只有这样组才会批准它们进行交互。如果一个成员的行为有损组的形象或者削弱了组的信誉,以后组内的其他成员将会拒绝与其合作,以此达到惩罚和警醒其他成员的作用。若一个节点长期有不好的行为,组内的成员将有权选举决定将其从组中除去。

每个节点需要交易时,总是率先考虑与自己同组的节点,其次是组信誉度的节点组中的节点。提供文件时,也总是优先为自己的同组节点提供服务,其次再考虑其他组成员。因此,游离在各个信誉组外的节点很难找到机会与其他节点交互,信誉值的提高也会很慢,被一个节点组接纳需要比较长的一段时间。但是如果恶意攻击其他节点或是更改其他节点的路由信息,信誉的下降幅度却会很快,而且还有可能被踢出组外。这样就可以起到一定的警示作用,使节点珍惜与组内其他节点的良好关系。

10.4 文件安全机制

文件安全机制是 P2P 分布自组织安全体系中的重要结构之一。文件的真实性问题和文件污染问题,得到了特别的关注,成为 P2P 诸多问题中的最为关注的问题之一[40]。

在无中心的 P2P 系统中,文件的真实性得不到保证,充斥着大量的虚假文件,甚至是威胁安全的恶意病毒或特洛伊木马,这将严重威胁终端实体,甚至影响整个网络的安全稳定运行。有研究表明,几乎所有的热门资源都有虚假文件,而且比想象的多得多。一些虚假文件的传播量甚至远大于真实文件,这不仅极大的浪费了网络带宽,给网络系统带来严重威胁,对于依靠大量用户支撑的 P2P 网络来说问题相当严重。

对于文件污染问题,是指在 P2P 文件共享系统中,恶意节点发布与指示主题不相符合的文件内容,并通过 P2P 文件共享进行传播。文件污染问题给 P2P 文件共享造成了很大的危害:首先,如果用户频繁遭遇污染文件,其感受到的可用 性会急剧降低,甚至最终放弃使用该系统;而且,它为病毒、蠕虫等恶意程序的传播提供了便利,造成了网络安全上的隐患。

对 P2P 网络的实际测量数据表明[41],现实存在的文件污染现象十分普遍,尤其是对于最近流行的内容。在 FastTrack/KaZaA、eDonkey、Overnet 等 P2P 系统中,有半数流行内容的副本是被污染的或是仿造的。

10.4.1 文件真实性概述

对等网络文件真实性是指:在对等网络中对于某请求节点的文件查询消息,会有不确

定数量的应答节点给予应答,确定哪些应答是满足条件的真实的文件,这就是确定对等网络文件的真实性。举例来说,如果一个节点发起一个"国富论"的查询,结果收到三个应答,这些应答中哪一个是真实的呢?应答之一可能恰好是亚当·斯密所著的《国富论》。另外的一个应答可能是修改了几个关键段落的亚当·斯密的《国富论》。第三个应答可能是某个网络写手对《国富论》的恶搞作品,该作品的名字也被命名为《国富论》。判断这些应答哪个是真实文件的过程就是文件真实性确定。

文件真实性是对等网络的一个安全要求,然而目前为止收到的关注并不是很多,Daswani 等提出过对等网络文件真实性认证的公开问题[42],他们在文献[42]中列举了四种不同的判断文件真实性的标准:

第一个判断准则是"最古老的文件是真实的"。这种定义认为最早提交到系统中的文档是真实副本。举例来说,如果亚当·斯密是《国富论》的作者并第一个向系统提交了这个文档,那么他的文档就被认为是"国富论"这一查询的真实文档。任何在此之后提交的名为"国富论"的文档都被认为是查询的非真实回应。利用时戳机制的方法可以帮助这种系统确立文件的真实性。然而可以看出来,这种机制往往太机械简单。

第二个判断的标准是以专家为基础确认文件真实性。在这方式中,一份文件签名如果经过了专家或权威节点的真实性确认,就被认为是真实的,例如,节点 A 是一个可信中心,他提供对签名的确认。当节点收到查询请求的回应后,他可以与这个可信中心 A 通信,用节点 A 中的记录来确认签名的真实性。其实这种方式是借鉴了 C/S 模式的集中管理原理,可信中心节点 A 实际上就是一个确认真实性的服务器。这种机制存在着相当大的局限,如果节点 A 暂时或永久失效、被攻击者入侵控制甚至 A 本身就是一个恶意节点,那么文件的真实性就根本得不到保障了。这是一个单点失效的问题。

第三个判断的标准是基于信誉确认文件真实性。在现实生活中,在一个领域有很多专家教授,但是他们对这个领域的造诣是有高低之分的,有的专家教授的观点更加权威可信,有的则次之。与在现实生活一样,一些专家节点可能比另一些专家节点更可信,可以在投票中加大可信专家选票的权重。这种权重要由一个完整的信誉机制来提供,这种机制要能保持、更新、传播信誉值。现在已经有一些信誉机制的研究,但至今仍没有一个成功地被商业界利用。

第四个判断文件真实性的准则是基于投票来确认文件真实性。为了解决前面第二个方案中信任中心节点 A 的单点失效问题,这种方案用投票的方法让许多专家对文件签名的真实性进行确认,只需一部分专家认同就认为文件是真实的。这就不再有单点失效的问题。

10.4.2 文件真实性确认协议

Ernesto Damiani 等人设计了 XREP 协议[38]是一种具有代表性意义的文件真实性确认协议,XREP 提出了结合文件信誉度(Reputation)和参与节点的信誉度来确定对等网络文件真实性的机制。在这种机制里,系统的每个参与节点拥有一个节点标识符 servent_id(一般是用该节点的公钥进行哈希计算得到的),每个文件拥有一个文件标识符 resource_id(一般就是把文件内容进行哈希计算得到的)。每个参与节点都有一个经验库(experience repository),经验库,记录对文件和其他通信节点的某些历史评价信息,如用二元组

(resource_id,value)记录文件的信誉度,value 以某种方法对文件的评价值,用三元组(servent_id,num_plus,num_minus)记录节点的信誉度。其中,num_plus 是在节点上成功下载文件的次数,num_minus 是在节点上文件下载不成功的次数,对节点的投票可以根据不同的标准,比如说一个简单的方法是,某节点只对成功下载次数为 0(num_plus=0)的其他节点给下面评价。整个搜索、投票和下载的过程具体可分为 5 个阶段:

1. 搜索资源

发起者 I 以类 Gnutella 形式的泛洪方式向所有邻居节点发送查询消息 Query,形如 Query(search_string,min_speed),系统中的节点在收到查询消息后立刻查看本地是否有符合查询请求的文件内容,如果有,就按照查询消息的发送路径返回一个查询响应消息,形如 QueryHit(num_hit,IP,port,speed,Result,trailer,servent_id),包括响应者节点标识符、查询到的文件名和其他信息组成的结果集 Result、匹配查询请求的文件数量、网速和<IP,port>对。

2. 选择资源和发起投票

根据上一步返回的查询响应结果,发起者 I 在收到的 QueryHit 中根据响应者的信息选择一个资源 r 和一定数量的资源 r 的提供者组成集合 $T=\{S_1, S_2, \cdots, S_n\}$,这种选择可以取决于请求者的个人喜好和同种资源提供者的人数。I 产生一对密钥对(PKpoll,SKpoll),然后将发起投票的消息 Poll(r,$\{S_1, S_2, \cdots, S_n\}$,PKpoll)以类 Gnutella 的方式泛洪出去,由此发起投票。系统中的节点收到发起投票消息以后,检查它们的经验库,根据对文件和节点的记录产生投票,投票用 I 的公钥加密按照投票发起消息的发送路径返回给请求者,形如 PollReply({IP,port,votes)},PKpoll)),用公钥加密有两个目的,其一是防止消息在传输过程中被人恶意窜改,其二是保证投票和投票者的机密性,不让攻击者发现投票者和票的关联。

3. 统计票数和核实投票

根据上一阶段收集的 PollReply,发起者 I 先用私钥 Skpoll 解密发现被篡改过的票并且将其丢弃,再根据 IP 地址排除派系,然后在排除派系后的所有投票者中选择一个投票者子集 V',再用<IP,port>向每一个投票者 vj(vj 在子集 V' 中)直接发起投票核实请求消息 TrueVote,要求 vj 向 I 发送核心信息 TrueVoteReply(response)核实投票,经过核实,I 相信某些投票,丢弃那些没有回复和没有通过核实的投票。

4. 选择资源提供者并检查其可用性

请求者根据从第 3 阶段中确定的可信投票选择一个信誉度最高的资源节点作为资源提供者 S。然而在资源下载之前,必须要核实资源提供者 S 的身份,因为要防止其他节点利用该资源提供者 S 的 servent_ids 冒充 S 提供资源下载。这个阶段的过程是:发起者 I 发送确认请求 AreYou(server_ids,r)给资源提供者节点,要求该节点作出应答,收到确认请求的节点用自己的私钥 SKs 加密确认消息后,连同自己的公钥 PKs 一同发送给请求者,形如 AreYouReply([response]SKs,PKs),请求者收到 AreYouReply 后用 PKs 解密[response]SKs,得到确认结果,然后把 PKs 作哈希计算,如果哈希计算的结果是 server_ids,则证明发起者 I 是在和真正的资源提供者 S 通信。

5. 下载资源

发起者 I 直接与资源提供者 S 通信，download(r)，要求下载资源，下载后请求者会计算文件内容的摘要以确定完整性，最后发起者 I 更新经验库中资源和资源提供者的记录。

由以上对 5 个阶段的分析可知，整个机制可以防止攻击者修改投票，能够发现和排除派系选票，还能有效地挫败攻击者假冒资源提供者。像 Gnutella 网络一样，系统具有幂规律(PowerLaw)特性，即热门一些的文件会比不热门的文件更频繁的被搜索到，对它们的投票也会多一些，而且少数节点有较高的度，多数节点的度较低，因此少数节点将提供多数的资源下载。

10.4.3 P2P 文件污染概述

所谓文件污染是指 P2P 文件共享网络中的恶意用户，可称之为“污染者”，将虚假甚至含有恶意内容的文件贴上某些热门内容的标签进行发布，诱骗其他用户下载，并利用 P2P 网络的自由共享功能进行更广泛散播的现象。

P2P 文件污染的危害主要有三方面：一是降低了网络内共享资源的可用性，二是破坏了安全和互利的共享资源环境，三是为病毒、蠕虫的传播提供了便利。由于现有的 P2P 文件共享网络普遍缺乏准入控制和内容管理机制，所以，污染者可以像正常用户一样自由地发布和共享任何内容。再加上 P2P 网络中文件传播往往是“一传十、十传百”，而用户的行为又具有很强的自主性，很难加以管理，因此，文件污染一旦发生，将很难得到有效的控制。

P2P 文件污染最初是版权组织为了破坏版权文件在 P2P 网络上的非法传播而采取的一种比较消极的技术手段。从 2002 年起，一些公司便开始雇佣 P2P 污染者，在各 P2P 网上布满其试图保护的音乐、电影、软件的假冒伪劣版本，欲使这些网络瘫痪。例如当时最著名的专业 P2P 污染公司 Overpeer，就曾经于 2003 年成功地使当时最受欢迎的 KaZaA/FastTrack 网络上被污染的文件占到总文件数量的一半以上。虽然从这些年的数据来看，这一技术并未真正达到版权保护的作用，但它对 P2P 文件共享系统造成的影响却吹响了 P2P 系统中内容安全对抗战的号角，并必将引发更大范围内不同目的性的文件污染和对抗以及更多的内容安全问题。因此，在现阶段分析讨论和防范文件污染不仅仅是个有关版权的议题，更是应对未来 P2P 系统中将要不断涌现的内容安全问题的一种必要的准备。

10.4.4 P2P 文件污染的研究

1. 文件污染方式

文件污染一般是针对某些选定的关键词进行的。文件污染者有如下三种污染方式可以选择[43]：

1）索引污染

最简单的文件污染方式是“索引污染”，即：在 P2P 网络的索引服务系统中注入大量虚假的记录，这些记录指向不存在的版本和/或副本。当用户按照这些记录的指示尝试下载时，将得到“无法链接”的提示。如果注入的虚假索引记录足够多，那么没有耐心的用户可能在几次失败的尝试之后放弃下载的努力。

索引污染既可以针对版本也可以针对副本。它与普通的版本污染和副本污染的不同之处在于，污染者注入网络中的索引记录指向并不存在的对象，因此污染者并不需要拥有强大的污染服务器来提供大量的上传服务。

2）副本污染

所谓副本污染，指的是污染者声称自己存有某个正确版本的副本，但实际上传给下载者的却是错误的数据。如果这种污染者足够多，那么，即使下载者能够挑选出正确的版本，一旦误选这些污染者作为下载源，也会浪费大量的时间精力和网络带宽。

这种污染方式要求污染者拥有强大的污染服务器来提供大量的上传服务。

3）版本污染

最复杂、也是危害性最大的文件污染方式是版本污染。实施版本污染的污染者首先针对一个（或同时针对多个）目标关键词制造出大量含有恶意或错误内容的污染版本。然后污染者将这些版本的索引信息注入目标 P2P 网络，并在其污染服务器上提供大量可供下载的副本。如果没有有效的识别措施和管理机制，网络中的用户在搜索相关主题时就很容易被这些具有大量可下载副本的污染版本所吸引。一旦下载了污染版本而又没有及时加以检验，一般用户很可能将该版本的本地副本设置为共享，并提供给其他用户下载。如此一来，污染版本将在网络中广泛的传播开来，甚至会超过了正确版本的副本数量，最终将正确副本淹没在污染副本中，使得该主题资源变得不可用。

P2P 共享文件的污染版本有很多不同的表现形式，例如，对于 MP3 歌曲文件，污染者可以采用截短、插入噪声、插入不可解码的数据片断甚至插入辱骂词句等方式来制造污染版本，而对于可执行文件，则可能是插入蠕虫、木马等恶意代码。由于 P2P 网络中共享资源的多样性，对文件版本的好坏，很难有效的自动识别措施，因此，版本污染具有很强的隐蔽性，大多数情况下只能依靠人工的识别。正是这种人工识别的滞后性，使得 P2P 网络中被污染的文件版本不仅可以通过污染服务器直接散发，还可以通过正常用户的共享行为得到更加广泛和迅速的传播。

以上三种污染方式可能单独使用，也可能结合起来形成更为复杂和隐蔽的复合式污染方式。需要指出的是，索引污染只会降低资源可用性而不会引入有害内容，副本污染虽然可能引入有害内容，但很容易通过校验码等简单的机制加以识别（目前的 P2P 文件共享网络大都采用了这种机制），因此这两种污染方式单独使用时危害性相对较小。对 P2P 共享社区的内容安全威胁最大的是版本污染，这种污染方式需要特别的关注，以下除非特别指出，提到“文件污染”的地方都是特指版本污染。

由于现有的 P2P 文件共享网络普遍缺乏准入控制和内容管理机制，所以，污染者可以像正常用户一样自由地发布和共享任何内容。再加上 P2P 网络中文件传播往往是“一传十、十传百”，而用户的行为又具有很强的自主性，很难加以管理，因此，文件污染一旦发生，一般很难得到有效的控制。最终的后果是造成 P2P 网络中充斥大量不可用的、甚至是有害的索引信息、文件版本或文件副本，从而引起有害内容的扩散和资源可用性的降低，并最终造成用户的流失。

2. 研究现状

P2P 文件污染的报道最早出现于 2002 年，美国的 Overpeer 公司曾经成功的使当时最

受欢迎的 KaZaA 网络上被污染的文件占到总文件数量的一半以上[44]。学术界对这一现象的研究是从 2005 年开始的。文献[41]最先对 P2P 文件污染进行了正式的描述，随后，文件污染问题开始引起更多研究者的关注。根据关注角度和研究方法的不同，目前学术界对 P2P 文件污染的研究工作大致可以分为三类：

第一类是对现有 P2P 网络中的文件污染现状的测量分析。例如文献[41]对 Kazaa 网络中的文件污染进行了详细的测量、统计和分析，文献[48]侧重于 eDonkey 网络，文献[45]侧重于 BitTorrent 网络，而文献[46]则同时对 eDonkey、KaZaa、Gnutella 和 OverNet 这几种流行的 P2P 网络上的内容可用性和污染问题进行了广泛深入的测量和分析。这些测量和分析很好的揭示了 P2P 网络中污染文件的静态空间分布，但大多并不能揭示出污染扩散过程的动态时间演化规律。

第二类是对 P2P 环境中文件污染问题的抽象和建模研究。文献[47]中的 P2P 文件传播模型借鉴了疾病传播的 K-M(Kermack-Mckendrick)模型。作者用他们的模型分析了单个版本文件的传播规律，并推导出保证传播成功的“离开率”门限。他们还通过仿真实验分析了两个版本(一个正常，一个被污染)的竞争传播。文献 [49]采用了类似的方法为 P2P 病毒和文件污染散播建模。他们还考察了节点动态进出系统的影响，以及信誉系统的作用。文献[50]中的模型较为简略，基本上遵循了相同的思路。但该文中对用户行为的问卷分析结果提供了非常有参考价值的两个结论，即：同一用户对不同类型的污染所具有的警觉性可能是不同的，而不同用户对同样的污染也可能具有不同的警觉性。

第三类研究工作的主要内容是探讨如何防治 P2P 文件污染或者对污染内容的扩散进行控制。由于 P2P 文件污染现象出现的历史短、规模大、情况复杂多变，所以这一类研究工作大多是尝试和探讨，并没有得到广泛公认的成功先例。被讨论较多的一种方案是针对共享文件的内容可用性而建立的“对象信誉系统”，关于这种方案的细节可以参见文献[51]。

10.5　小结

本章对对等网络的可信及公正性问题进行了研究，综述了当前研究中采用的激励机制、信任机制和安全机制来控制节点的自私行为和恶意行为。随着对等网络应用日益广泛，安全问题也将越来越突出，对对等网络的安全研究工作将更为系统和深入。

参考文献

[1] Douceur R J. The Sybil attack. In Proceeding for the 1st International Workshop on Peer-to-Peer Systems(IPTPS’02)，Cambridge，Massachusetts，March 2002.

[2] Srivatsa M，Liu L. Vulnerabilities and Security Threats in Structured Peer-to-Peer Systems：A Quantitative Analysis，Proceedings of the 20th Annual Computer Security Applications Conference，2004.

[3] Adar E，Huberman B A. Free Riding on Gnetella. First Monday，October 2000，5(10).

[4] Xiong L，Liu L. A Reputation-Base Trust Model for Peer-to-Peer e-Commerce Communities. In ACM Conference(ACSAC 2004)，Tucson，Arizona，December 2004：6-10.

[5] Saroiu S，Gummadi P K，Gribble S D. A measurement study of peer-to-peer le sharing systems. In

Proceedings of Multimedia Computing and Networking 2002 (MMCN),San Jose,CA,January 2002.
[6] Kollock P. The Economies of Online Cooperation: Gifts and Public Goods in Cyberspace. Communicaties in Cyberspace,Chapter 9,1999.
[7] Hardin,G. The Tragedy of the Commons. Science. 1968 Dec 13,162(5364): 1243-8.
[8] Feldman M,Lai K,Stoica I,et al. Robust Incentive Techniques for Peer-to-Peer Networks. ACM E-Commerce Conference (EC'04),May 2004.
[9] http://www. kazaa. com.
[10] Kung H T,Wu C H. Differentiated Admission for Peer-to-Peer Systems: Incentivizing Peers to Contribute Their Resources.
[11] Golle P,K Leyton-Brown,Mironov I. Incentives for Sharing in Peer-to-Peer Networks. In ACM Conference on Electronic Commerce,2002.
[12] Yang B,Garcia-Molina H. PPay: Micropayments for Peer-to-Peer Systems. In Proceedings of the 10th ACM Conference on Computer and Communications Security(CCS),OCT 2003.
[13] Ioannidis J,Ioannidis S,et al. Fileteller: Paying and Getting Paid for File Storage,Springer-Verlag; 2003,1999: 282-300.
[14] Vishnumurthy V,Chandrakumar S,et al. Karma: A Secure Economic Framework for P2P Resource sharing. 2003.
[15] Hausheer D, Stiller B. PeerMint: Decentralized and Secure Accounting for Peer-to-Peer Applications,Springer. 2005.
[16] MojoNation,http://www. mojonation. net.
[17] Dabek F,Kaashoek M F,et al. Wide-Area Cooperative Storage with CFS,ACM New York,NY, 2001,35: 202-215.
[18] Adya A, Bolosky W J, et al. FARSITE: Federated, Available, and Reliable Storage for an Incompletely Trusted Environment,ACM Association for Computing Machinery. 2002,36: 1-14.
[19] Cox L P,Murray C D,et al. Pastiche: making backup cheap and easy,ACM New York,NY,2002, 36: 285-298.
[20] Cox L P,Noble B D. Samsara: honor among thieves in Peer-to-Peer storage,ACM New York,NY, 2003: 120-132.
[21] Sun Q,Garcia-Molina H. SLIC: A Selfish Link-Based Incentive Mechanism for Unstructured Peer-to-Peer Networks,2004: 506-515.
[22] Axelrod R,Hamilton W D. The evolution of cooperation. 1981,211: 1390-1396.
[23] Cohen B. Incentives Build Robustness in BitTorrent,Berkeley,CA,2003,6.
[24] Riolo R L,Cohen M D,et al. Evolution of Cooperation Without Reciprocity. 2001,414: 441-443.
[25] Hales D,Edmonds B. Applying a Socially-Inspired Technique (tags) to Improve Cooperation in P2P Networks. IEEE Transactions in Systems, Man and Cybernetics-Part A: Systems and Humans, 2005,35(3): 385-395.
[26] Kamvar S D,Schlosser M T,et al. The Eigen Trust Algorithm for Reputation Management in P2P networks. New York: ACM Press,2003: 640-651.
[27] Luhmann N. Trust and Power. Chichester: Wiley,1979.
[28] Gambetta D. Can we trust trust? Gambetta D,ed. In: Trust: Making and Breaking Cooperative Relations. Basil Blackwell: Oxford Press,1990: 213～237.
[29] Randison T,Sloman M. A Survey of Trust in Internet Applications. 2000,3: 2-16.
[30] Chen R,Yeager W. Poblano: A distributed trust model for P2P networks. Technical Report,TR-I4-02-08,Palo Alto: Sun Microsystem,2002.
[31] Alfarez Abdul-Rahman,Stephen Hailes. Supporting Trust in Virtual Communities. Proceedings of

the 33rd Hawaii International Conference on System Sciences,January 2000,6(6007)：04-07.
[32] Garfinkel S. Pretty Good Privacy. O'Reilly,1995.
[33] Cornelli F,Damiani E,et al. Choosing Reputable Servents in a P2P Network,ACM New York,NY,2002：376-386.
[34] Yao Wang, Julita Vasileva. Trust and Reputation Model in Peer-to-Peer Networks. P2P'03,2003.
[35] Wang Yao,Vassileva Julila. Bayesian Network Model in Peer-to-Peer Networks Agents and Peer-to-Peer. Computing AP2PC 2003,2003：23-34.
[36] Josang A. Trust-Based Decision Making for Electronic Transactions. Proceedings of the Fourth NoSDic workshop on Secure Computer Systems (NOSDSEC'99). Stockholm University, Sweden,1999.
[37] Abdul Rahmma A, Hailes S. Supporting Trust in Virtual Communities In Proceedings Hawaii International Conference on System Sciences 33,Maui,Hawaii,January 2000：4-7.
[38] Damiani E, D. C. di Vimereati, S. Paraboschi, et al, A Reputation-Based Approach for Choosing Reliable Resources in Peer-to-Peer Networks. In Proc. of the 9th ACM Conference On Computer and Communications Security. 2002.
[39] Kamvar S D, Schlossc M. EigenRep: Reputation Management in P2P Networks. The 12th International World Wide Web ConferenCe. Budapest,Hungary：ACM press. 2003.
[40] Christin N,Weigend A S,et al. Content availability,pollution and poisoning in file sharing Peer-to-Peer Networks,ACM New York,2005,68-77.
[41] Liang J, Kumar R,et al. (2005). Pollution in P2P file sharing systems.
[42] Daswani N,Garcia Molina H,et al. Open Problems in Data-Sharing Peer-to-Peer Systems,Springer,2003：1-15.
[43] 左敏,李建华. P2P 中的文件污染与污染防治,计算机工程,2007,38(18).
[44] Overpeer Spreads Fake Files Through P2P Networks,http://www. afterdawn. com/news/archive/3101. cfm.
[45] Pouwelse J,Garbacki P,et al. The Bittorrent P2P File-Sharing System：Measurements and Analysis,Springer. 2005,3640：205.
[46] Christin N,Weigend A S,et al. Content availability,pollution and poisoning in file sharing Peer-to-Peer Networks,ACM New York,2005：68-77.
[47] Gu Q, Bai K,et al. (2006). Modeling of pollution in P2P file sharing systems.
[48] Leibnitz K,Tobias Hoβfeld,Wakamiya N,et al. On Pollution in eDonkey-like Peer-to-Peer File-Sharing Networks,13th GI/ITG Conference on Measurement,Modeling,and Evaluation of Computer and Communication Systems (MMB 2006),Nürnberg,Germany,March 2006.
[49] Thommes R,Coates M. Epidemiological Models of Peer-to-Peer Viruses and Pollution. Technical report,Mc Gill University,Department of Electrical and Computer Engineerning,June 2005.
[50] Lee U,Choi M,et al. Understanding Pollution Dynamics in P2P File Sharing. IPTPS06,2005.
[51] Walsh K,Sirer E G. Thwarting P2P Pollution Using Object Reputation,Tech Rep Computer Science Department Technical Report TR 2005-1980,Cornell University,Feb 2005.
[52] 窦文,王怀民,贾焰等. 构造基于推荐的 Peer-to-Peer 环境下的 Trust 模型. 软件学报,2004,15(4)：571-583.
[53] 宋雪吕,陆建德. 基于组群的 P2P 系统中的信誉机制. 微机发展,2005,15(11)：11-13.
[54] Gummadi A,Yoon J P. Modeling Group Trust for Peer-to-Peer Access Control. In Proceedings of the 15th International Workshop on Database and Expert Systems Applications (DEXA'04). 2004.

读者意见反馈

亲爱的读者：

感谢您一直以来对清华版计算机教材的支持和爱护。为了今后为您提供更优秀的教材，请您抽出宝贵的时间来填写下面的意见反馈表，以便我们更好地对本教材做进一步改进。同时如果您在使用本教材的过程中遇到了什么问题，或者有什么好的建议，也请您来信告诉我们。

地址：北京市海淀区双清路学研大厦 A 座 602 室 计算机与信息分社营销室 收

邮编：100084　　电子邮箱：jsjjc@tup.tsinghua.edu.cn

电话：010-62770175-4608/4409　　邮购电话：010-62786544

教材名称：无线自组织网络和对等网络原理与安全

ISBN 978-7-302-19933-5

个人资料

姓名：________ 年龄：______ 所在院校/专业：__________

文化程度：______ 通信地址：______________

联系电话：______ 电子信箱：______________

您使用本书是作为：□指定教材 □选用教材 □辅导教材 □自学教材

您对本书封面设计的满意度：

□很满意 □满意 □一般 □不满意　改进建议____________

您对本书印刷质量的满意度：

□很满意 □满意 □一般 □不满意　改进建议____________

您对本书的总体满意度：

从语言质量角度看　□很满意 □满意 □一般 □不满意

从科技含量角度看　□很满意 □满意 □一般 □不满意

本书最令您满意的是：

□指导明确 □内容充实 □讲解详尽 □实例丰富

您认为本书在哪些地方应进行修改？（可附页）

____

____

您希望本书在哪些方面进行改进？（可附页）

____

____

电子教案支持

敬爱的教师：

为了配合本课程的教学需要，本教材配有配套的电子教案（素材），有需求的教师可以与我们联系，我们将向使用本教材进行教学的教师免费赠送电子教案（素材），希望有助于教学活动的开展。相关信息请拨打电话 010-62776969 或发送电子邮件至 jsjjc@tup.tsinghua.edu.cn 咨询，也可以到清华大学出版社主页（http://www.tup.com.cn 或 http://www.tup.tsinghua.edu.cn）上查询。